



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07H 21/02, C07K 1/00, A61K 39/02, A01N 37/18		A1	(11) International Publication Number: WO 98/21225
			(43) International Publication Date: 22 May 1998 (22.05.98)
(21) International Application Number: PCT/US97/21353		(72) Inventors: HAAS, Rainer; Ursrainer Ring 65, D-72076 Tuebingen (DE). KLEANTHOS, Harold; 89 Madison Avenue, Newtonville, MA 02160 (US). TOMB, Jean-Francois; 3501 St. Paul Street, Baltimore, MD 21202 (US). MILLER, Charles; 32 Maple Avenue, Medford, MA 02155 (US). AL-GARAWI, Amal; 32 Garrison Street #4501, Boston, MA 02116 (US). ODENBREIT, Stefan; Mazartstrasse 8, D-72119 Ammerbuch (DE). MEYER, Thomas; Spe-mannstrasse 30, D-72076 Tuebingen (DE).	
(22) International Filing Date: 14 November 1997 (14.11.97)		(74) Agent: CLARK, Paul, T.; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110-2214 (US).	
(30) Priority Data:			
08/749,051	14 November 1996 (14.11.96)	US	
08/831,309	1 April 1997 (01.04.97)	US	
08/834,705	1 April 1997 (01.04.97)	US	
08/833,457	1 April 1997 (01.04.97)	US	
08/881,227	24 June 1997 (24.06.97)	US	
08/902,615	29 July 1997 (29.07.97)	US	
(71) Applicants: MERIEUX ORAVAX SOCIETE EN NOM COL-LECTIF PASTEUR MERIEUX SERUMS ET VACCINS S.A. [FR/FR]; 58, Avenue Leclerc, F-69007 Lyon (FR). MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. BERLIN [DE/DE]; D-80539 München (DE). HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: HELICOBACTER POLYPEPTIDES AND CORRESPONDING POLYNUCLEOTIDE MOLECULES			
(57) Abstract			
<p>The invention provides <i>Helicobacter</i> polypeptides that can be used in vaccination methods for preventing or treating <i>Helicobacter</i> infection, and polynucleotides that encode these polypeptides.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HELICOBACTER POLYPEPTIDES AND CORRESPONDING
POLYNUCLEOTIDE MOLECULES

5 The invention relates to *Helicobacter* antigens and corresponding polynucleotide molecules that can be used in methods to prevent or treat *Helicobacter* infection in mammals, such as humans.

Background of the Invention

Helicobacter is a genus of spiral, gram-negative bacteria that
10 colonize the gastrointestinal tracts of mammals. Several species colonize the stomach, most notably *H. pylori*, *H. heilmanii*, *H. felis*, and *H. mustelae*. Although *H. pylori* is the species most commonly associated with human infection, *H. heilmanii* and *H. felis* have also been isolated from humans, but at lower frequencies than *H. pylori*. *Helicobacter* infects over 50% of adult
15 populations in developed countries and nearly 100% in developing countries and some Pacific rim countries, making it one of the most prevalent infections worldwide.

Helicobacter is routinely recovered from gastric biopsies of humans with histological evidence of gastritis and peptic ulceration. Indeed, *H. pylori*
20 is now recognized as an important pathogen of humans, in that the chronic gastritis it causes is a risk factor for the development of peptic ulcer diseases and gastric carcinoma. It is thus highly desirable to develop safe and effective vaccines for preventing and treating *Helicobacter* infection.

 A number of *Helicobacter* antigens have been characterized or
25 isolated. These include urease, which is composed of two structural subunits of approximately 30 and 67 kDa (Hu *et al.*, Infect. Immun. 58:992, 1990; Dunn *et*

-2-

al., J. Biol. Chem. 265:9464, 1990; Evans *et al.*, Microbial Pathogenesis 10:15, 1991; Labigne *et al.*, J. Bact., 173:1920, 1991); the 87 kDa vacuolar cytotoxin (VacA) (Cover *et al.*, J. Biol. Chem. 267:10570, 1992; Phadnis *et al.*, Infect. Immun. 62:1557, 1994; WO 93/18150); a 128 kDa immunodominant antigen associated with the cytotoxin (CagA, also called TagA; WO 93/18150; U.S. Patent No. 5,403,924); 13 and 58 kDa heat shock proteins HspA and HspB (Suerbaum *et al.*, Mol. Microbiol. 14:959, 1994; WO 93/18150); a 54 kDa catalase (Hazell *et al.*, J. Gen. Microbiol. 137:57, 1991); a 15 kDa histidine-rich protein (Hpn) (Gilbert *et al.*, Infect. Immun. 63:2682, 1995); a 20 kDa membrane-associated lipoprotein (Kostreynska *et al.*, J. Bact. 176:5938, 1994); a 30 kDa outer membrane protein (Bölin *et al.*, J. Clin. Microbiol. 33:381, 1995); a lactoferrin receptor (FR 2,724,936); and several porins, designated HopA, HopB, HopC, HopD, and HopE, which have molecular weights of 48-67 kDa (Exner *et al.*, Infect. Immun. 63:1567, 1995; Doig *et al.*, J. Bact. 177:5447, 1995). Some of these proteins have been proposed as potential vaccine antigens. In particular, urease is believed to be a vaccine candidate (WO 94/9823; WO 95/22987; WO 95/3824; Michetti *et al.*, Gastroenterology 107:1002, 1994). Nevertheless, it is thought that several antigens may ultimately be necessary in a vaccine.

20

Summary of the Invention

The invention provides polynucleotide molecules that encode *Helicobacter* polypeptides, designated GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596, GHPO 699, GHPO 724, GHPO 730, GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100,

-3-

GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559, GHPO 427, GHPO 1045, GHPO 1262, GHPO 1688, GHPO 1538, GHPO 346, GHPO 1012, GHPO 470, GHPO 1398, GHPO 1550, GHPO 276, GHPO 1501, GHPO 706, GHPO 1001, GHPO 732, GHPO 329, GHPO 574, GHPO 1190, GHPO 1374, GHPO 1620, GHPO 956, HPO 98, GHPO 689, GHPO 208, GHPO 296, GHPO 726, GHPO 1026, GHPO 1301, GHPO 1536, GHPO 166, GHPO 253, GHPO 297, GHPO 615, GHPO 1278, GHPO 1282, GHPO 1420, GHPO 1484, GHPO 1719, and GHPO 1252, which can be used, *e.g.*, in methods to prevent, treat, or diagnose *Helicobacter* infection. The polypeptides of the invention include those having the amino acid sequences shown in SEQ ID NOs:2-170 (even numbers), as well as mature forms of proteins having sequences shown in SEQ ID NOs:2-170 in their unprocessed forms, and fragments thereof. Those skilled in the art will understand that the invention also includes polynucleotide molecules that encode mutants and derivatives of these polypeptides, which can result from the addition, deletion, or substitution of non-essential amino acids, as is described further below.

In addition to the polynucleotide molecules described above, the invention includes the corresponding polypeptides (*i.e.*, polypeptides encoded by the polynucleotide molecules of the invention, or fragments thereof), and monospecific antibodies that specifically bind to these polypeptides.

The present invention has many applications and includes expression cassettes, vectors, and cells transformed or transfected with the polynucleotides of the invention. Accordingly, the present invention provides (i) methods for producing polypeptides of the invention in recombinant host systems and

-4-

related expression cassettes, vectors, and transformed or transfected cells; (ii) live vaccine vectors, such as pox virus, *Salmonella typhimurium*, and *Vibrio cholerae* vectors, that contain polynucleotides of the invention (such vaccine vectors being useful in, e.g., methods for preventing or treating *Helicobacter* infection) in combination with a diluent or carrier, and related pharmaceutical compositions and associated therapeutic and/or prophylactic methods; (iii) therapeutic and/or prophylactic methods involving administration of polynucleotide molecules, either in a naked form or formulated with a delivery vehicle, polypeptides or mixtures of polypeptides, or monospecific antibodies of the invention, and related pharmaceutical compositions; (iv) methods for detecting the presence of *Helicobacter* in biological samples, which can involve the use of polynucleotide molecules, monospecific antibodies, or polypeptides of the invention; and (v) methods for purifying polypeptides of the invention by antibody-based affinity chromatography.

Brief Description of the Drawings

Fig. 1A is a diagrammatic representation of transposon TnMax9, which is a derivative of the TnMax transposon system (Haas *et al.*, Gene 130:23-21, 1993). The mini-transposon carries the *blaM* gene, which is the β -lactamase gene lacking a promoter and a signal sequence, next to the inverted repeats (IR) and the M13 forward (M13-FP) and reverse (M13-RP1) primer binding sites. The resolution site (*res*) and an origin of replication (*ori_{rd}*) are located between the *BlaM* gene and the constitutive *cat_{GC}*-resistance gene. The transposase *tnpA* and resolvase *tnpR* genes are located outside of the mini-transposon and are under the control of the inducible P_{uc} promoter. The *lacIq* gene encodes the Lac repressor.

-5-

Fig. 1B is a diagrammatic representation of plasmid pMin2. pMin2 contains a multiple cloning site, the tetracycline resistance gene (*tet*), an origin of transfer (*oriT*), an origin of replication (*ori_{ColE1}*), a transcriptional terminator (*t_{td}*), and a weak, constitutive promoter (*P_{iga}*). *H. pylori* chromosome fragments were introduced into the *Bgl*III and *Cla*I sites of pMin2.

Detailed Description

Open reading frames (ORFs) encoding new, full length polypeptides, designated GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596, GHPO 699, GHPO 724, GHPO 730, GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100, GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559, GHPO 427, GHPO 1045, GHPO 1262, GHPO 1688, GHPO 1538, GHPO 346, GHPO 1012, GHPO 470, GHPO 1398, GHPO 1550, GHPO 276, GHPO 1501, GHPO 706, GHPO 1001, GHPO 732, GHPO 329, GHPO 574, GHPO 1190, GHPO 1374, GHPO 1620, GHPO 956, GHPO 98, GHPO 689, GHPO 208, GHPO 296, GHPO 726, GHPO 1026, GHPO 1301, GHPO 1536, GHPO 166, GHPO 253, GHPO 297, GHPO 615, GHPO 1278, GHPO 1282, GHPO 1420, GHPO 1484, GHPO 1719, and GHPO 1252, have been identified in the *H. pylori* genome. These polypeptides can be used, for example, in vaccination methods for preventing or treating *Helicobacter* infection. Some of the new polypeptides are secreted polypeptides that can be produced in their mature forms (*i.e.*, as polypeptides that have been exported through class II or class III

-6-

secretion pathways) or as precursors that include signal peptides, which can be removed in the course of excretion/secretion by cleavage at the N-terminal end of the mature form. (The cleavage site is located at the C-terminal end of the signal peptide, adjacent to the mature form.)

5 According to a first aspect of the invention, there are provided isolated polynucleotides that encode the precursor and mature forms of *Helicobacter* GHPO 13, GHPO 73, GHPO 90, GHPO 107, GHPO 136, GHPO 191, GHPO 213, GHPO 240, GHPO 408, GHPO 411, GHPO 419, GHPO 431, GHPO 474, GHPO 591, GHPO 596, GHPO 699, GHPO 724, GHPO 730,
10 GHPO 761, GHPO 804, GHPO 805, GHPO 812, GHPO 879, GHPO 888, GHPO 986, GHPO 1056, GHPO 1081, GHPO 1100, GHPO 1140, GHPO 1148, GHPO 1200, GHPO 1212, GHPO 1258, GHPO 1263, GHPO 1273, GHPO 1284, GHPO 1299, GHPO 1327, GHPO 1346, GHPO 1378, GHPO 1412, GHPO 1443, GHPO 1466, GHPO 1476, GHPO 1536, GHPO 1559,
15 GHPO 427, GHPO 1045, GHPO 1262, GHPO 1688, GHPO 1538, GHPO 346, GHPO 1012, GHPO 470, GHPO 1398, GHPO 1550, GHPO 276, GHPO 1501, GHPO 706, GHPO 1001, GHPO 732, GHPO 329, GHPO 574, GHPO 1190, GHPO 1374, GHPO 1620, GHPO 956, HPO 98, GHPO 689, GHPO 208, GHPO 296, GHPO 726, GHPO 1026, GHPO 1301, GHPO 1536, GHPO 166,
20 GHPO 253, GHPO 297, GHPO 615, GHPO 1278, GHPO 1282, GHPO 1420, GHPO 1484, GHPO 1719, and GHPO 1252.

An isolated polynucleotide of the invention encodes:

(i) a polypeptide having an amino acid sequence that is homologous to a *Helicobacter* amino acid sequence of a polypeptide, the *Helicobacter*
25 amino acid sequence being selected from the group consisting of the amino acid sequences shown in SEQ ID NO:2 (GHPO 13), SEQ ID NO:4 (GHPO 73), SEQ ID NO:6 (GHPO 90), SEQ ID NO:8 (GHPO 107), SEQ ID NO:10

-7-

(GHPO 136), SEQ ID NO:12 (GHPO 191), SEQ ID NO:14 (GHPO 213), SEQ ID NO:16 (GHPO 240), SEQ ID NO:18 (GHPO 408), SEQ ID NO:20 (GHPO 411), SEQ ID NO:22 (GHPO 419), SEQ ID NO:24 (GHPO 431), SEQ ID NO:26 (GHPO 474), SEQ ID NO:28 (GHPO 591), SEQ ID NO:30 (GHPO 596), SEQ ID NO:32 (GHPO 699), SEQ ID NO:34 (GHPO 724), SEQ ID NO:36 (GHPO 730), SEQ ID NO:38 (GHPO 761), SEQ ID NO:40 (GHPO 804), SEQ ID NO:42 (GHPO 805), SEQ ID NO:44 (GHPO 812), SEQ ID NO:46 (GHPO 879), SEQ ID NO:48 (GHPO 888), SEQ ID NO:50 (GHPO 986), SEQ ID NO:52 (GHPO 1056), SEQ ID NO:54 (GHPO 1081), SEQ ID NO:56 (GHPO 1100), SEQ ID NO:58 (GHPO 1140), SEQ ID NO:60 (GHPO 1148), SEQ ID NO:62 (GHPO 1200), SEQ ID NO:64 (GHPO 1212), SEQ ID NO:66 (GHPO 1258), SEQ ID NO:68 (GHPO 1263), SEQ ID NO:70 (GHPO 1273), SEQ ID NO:72 (GHPO 1284), SEQ ID NO:74 (GHPO 1299), SEQ ID NO:76 (GHPO 1327), SEQ ID NO:78 (GHPO 1346), SEQ ID NO:80 (GHPO 1378), SEQ ID NO:82 (GHPO 1412), SEQ ID NO:84 (GHPO 1443), SEQ ID NO:86 (GHPO 1466), SEQ ID NO:88 (GHPO 1476), SEQ ID NO:90 (GHPO 1536), SEQ ID NO:92 (GHPO 1559), SEQ ID NO:94 (GHPO 427), SEQ ID NO:96 (GHPO 1045), SEQ ID NO:98 (GHPO 1262), SEQ ID NO:100 (GHPO 1688), SEQ ID NO:102 (GHPO 1538), SEQ ID NO:104 (GHPO 346), SEQ ID NO:106 (GHPO 1012), SEQ ID NO:108 (GHPO 470), SEQ ID NO:110 (GHPO 1398), SEQ ID NO:112 (GHPO 1550), SEQ ID NO:114 (GHPO 276), SEQ ID NO:116 (GHPO 1501), SEQ ID NO:118 (GHPO 706), SEQ ID NO:120 (GHPO 1001), SEQ ID NO:122 (GHPO 732), SEQ ID NO:124 (GHPO 329), SEQ ID NO:126 (GHPO 574), SEQ ID NO:128 (GHPO 1190), SEQ ID NO:130 (GHPO 1374), SEQ ID NO:132 (GHPO 1620), SEQ ID NO:134 (GHPO 956), SEQ ID NO:136 (HPO 98), SEQ ID NO:138 (GHPO 689), SEQ ID NO:140 (GHPO 208), SEQ ID NO:142 (GHPO 296), SEQ ID

-8-

NO:144 (GHPO 726), SEQ ID NO:146 (GHPO 1026), SEQ ID NO:148 (GHPO 1301), SEQ ID NO:150 (GHPO 1536), SEQ ID NO:152 (GHPO 166), SEQ ID NO:154 (GHPO 253), SEQ ID NO:156 (GHPO 297), SEQ ID NO:158 (GHPO 615), SEQ ID NO:160 (GHPO 1278), SEQ ID NO:162 (GHPO 1282),
5 SEQ ID NO:164 (GHPO 1420), SEQ ID NO:166 (GHPO 1484), SEQ ID NO:168 (GHPO 1719), and SEQ ID NO:170 (GHPO 1252); or

(ii) a derivative of the polypeptide.

In addition to the full-length polypeptides encoded by the polynucleotides of the invention, as set forth above, polynucleotides included in
10 the invention can also encode polypeptides that lack signal sequences, as well as other polypeptide or peptide fragments of the full-length polypeptides.

The term "isolated polynucleotide" is defined as a polynucleotide that is removed from the environment in which it naturally occurs. For example, a naturally-occurring DNA molecule present in the genome of a
15 living bacteria or as part of a gene bank is not isolated, but the same molecule, separated from the remaining part of the bacterial genome, as a result of, *e.g.*, a cloning event (amplification), is "isolated." Typically, an isolated DNA molecule is free from DNA regions (*e.g.*, coding regions) with which it is immediately contiguous, at the 5' or 3' ends, in the naturally occurring genome.
20 Such isolated polynucleotides can be part of a vector or a composition and still be isolated, as such a vector or composition is not part of its natural environment.

A polynucleotide of the invention can consist of RNA or DNA (*e.g.*, cDNA, genomic DNA, or synthetic DNA), or modifications or combinations of
25 RNA or DNA. The polynucleotide can be double-stranded or single-stranded and, if single-stranded, can be the coding (sense) strand or the non-coding (anti-sense) strand. The sequences that encode polypeptides of the invention, as

shown in any of SEQ ID NOs:2-170 (even numbers), can be (a) the coding sequence as shown in any of SEQ ID NOs:1-169 (odd numbers); (b) a ribonucleotide sequence derived by transcription of (a); or (c) a different coding sequence that, as a result of the redundancy or degeneracy of the genetic code, encodes the same polypeptides as the polynucleotide molecules having the sequences illustrated in any of SEQ ID NOs:1-169 (odd numbers). The polypeptide can be one that is naturally secreted or excreted by, *e.g.*, *H. felis*, *H. mustelae*, *H. heilmanii*, or *H. pylori*.

By "polypeptide" or "protein" is meant any chain of amino acids, regardless of length or post-translational modification (*e.g.*, glycosylation or phosphorylation). Both terms are used interchangeably in the present application.

By "homologous amino acid sequence" is meant an amino acid sequence that differs from an amino acid sequence shown in any of SEQ ID NOs:2-170 (even numbers), or an amino acid sequence encoded by the nucleotide sequence of any of SEQ ID NOs:1-169 (odd numbers), by one or more non-conservative amino acid substitutions, deletions, or additions located at positions at which they do not destroy the specific antigenicity of the polypeptide. Preferably, such a sequence is at least 75%, more preferably at least 80%, and most preferably at least 90% identical to an amino acid sequence shown in any of SEQ ID NOs:2-170 (even numbers).

Homologous amino acid sequences include sequences that are identical or substantially identical to an amino acid sequence as shown in any of SEQ ID NOs:2-170 (even numbers). By "amino acid sequence that is substantially identical" is meant a sequence that is at least 90%, preferably at least 95%, more preferably at least 97%, and most preferably at least 99% identical to an amino acid sequence of reference and that differs from the

-10-

sequence of reference, if at all, by a majority of conservative amino acid substitutions.

Conservative amino acid substitutions typically include substitutions among amino acids of the same class. These classes include, for example, amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; amino acids having basic side chains, such as lysine, arginine, and histidine; amino acids having acidic side chains, such as aspartic acid and glutamic acid; and amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Homology can be measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Similar amino acid sequences are aligned to obtain the maximum degree of homology (*i.e.*, identity). To this end, it may be necessary to artificially introduce gaps into the sequence. Once the optimal alignment has been set up, the degree of homology (*i.e.*, identity) is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions.

Homologous polynucleotide sequences are defined in a similar way. Preferably, a homologous sequence is one that is at least 45%, more preferably at least 60%, and most preferably at least 85% identical to a coding sequence of any of SEQ ID NOs:1-169 (odd numbers).

Polypeptides having a sequence homologous to any one of the sequences shown in SEQ ID NOs:2-170 (even numbers), include naturally-occurring allelic variants, as well as mutants or any other non-naturally

-11-

occurring variants that are analogous in terms of antigenicity, to a polypeptide having a sequence as shown in any one of SEQ ID NOs:2-170 (even numbers).

As is known in the art, an allelic variant is an alternate form of a polypeptide that is characterized as having a substitution, deletion, or addition of one or more amino acids that does not alter the biological function of the polypeptide. By "biological function" is meant a function of the polypeptide in the cells in which it naturally occurs, even if the function is not necessary for the growth or survival of the cells. For example, the biological function of a porin is to allow the entry into cells of compounds present in the extracellular medium. The biological function is distinct from the antigenic function. A polypeptide can have more than one biological function.

Allelic variants are very common in nature. For example, a bacterial species, *e.g.*, *H. pylori*, is usually represented by a variety of strains that differ from each other by minor allelic variations. Indeed, a polypeptide that fulfills the same biological function in different strains can have an amino acid sequence that is not identical in each of the strains. Such an allelic variation can be equally reflected at the polynucleotide level.

Support for the use of allelic variants of polypeptide antigens comes from, *e.g.*, studies of the *Helicobacter* urease antigen. The amino acid sequence of *Helicobacter* urease varies widely from species to species, yet cross-species protection occurs, indicating that the urease molecule, when used as an immunogen, is highly tolerant of amino acid variations. Even among different strains of the single species *H. pylori*, there are amino acid sequence variations.

For example, although the amino acid sequences of the UreA and UreB subunits of *H. pylori* and *H. felis* ureases differ from one another by 26.5% and 11.8%, respectively (Ferrero *et al.*, Molecular Microbiology

-12-

9(2):323-333, 1993), it has been shown that *H. pylori* urease protects mice from *H. felis* infection (Michetti *et al.*, Gastroenterology 107:1002, 1994). In addition, it has been shown that the individual structural subunits of urease, UreA and UreB, which contain distinct amino acid sequences, are both
5 protective antigens against *Helicobacter* infection (Michetti *et al.*, *supra*). Similarly, Cuenca *et al.* (Gastroenterology 110:1770, 1996) showed that therapeutic immunization of *H. mustelae*-infected ferrets with *H. pylori* urease was effective at eradicating *H. mustelae* infection. Further, several urease variants have been reported to be effective vaccine antigens, including, *e.g.*,
10 recombinant UreA + UreB apoenzyme expressed from pORV142 (UreA and UreB sequences derived from *H. pylori* strain CPM630; Lee *et al.*, J. Infect. Dis. 172:161, 1995); recombinant UreA + UreB apoenzyme expressed from pORV214 (UreA and UreB sequences differ from *H. pylori* strain CPM630 by one and two amino acid changes, respectively; Lee *et al.*, *supra*, 1995); a
15 UreA-glutathione-S-transferase fusion protein (UreA sequence from *H. pylori* strain ATCC 43504; Thomas *et al.*, Acta Gastro-Enterologica Belgica 56:54, 1993); UreA + UreB holoenzyme purified from *H. pylori* strain NCTC11637 (Marchetti *et al.*, Science 267:1655, 1995); a UreA-MBP fusion protein (UreA from *H. pylori* strain 85P; Ferrero *et al.*, Infection and Immunity 62:4981,
20 1994); a UreB-MBP fusion protein (UreB from *H. pylori* strain 85P; Ferrero *et al.*, *supra*); a UreA-MBP fusion protein (UreA from *H. felis* strain ATCC 49179; Ferrero *et al.*, *supra*); a UreB-MBP fusion protein (UreB from *H. felis* strain ATCC 49179; Ferrero *et al.*, *supra*); and a 37 kDa fragment of UreB containing amino acids 220-569 (Dore-Davin *et al.*, "A 37 kD fragment of
25 UreB is sufficient to confer protection against *Helicobacter felis* infection in mice"). Finally, Thomas *et al.* (*supra*) showed that oral immunization of mice

-13-

with crude sonicates of *H. pylori* protected mice from subsequent challenge with *H. felis*.

Polynucleotides, *e.g.*, DNA molecules, encoding allelic variants can easily be obtained by polymerase chain reaction (PCR) amplification of genomic bacterial DNA extracted by conventional methods. This involves the use of synthetic oligonucleotide primers matching sequences that are upstream and downstream of the 5' and 3' ends of the coding region. Suitable primers can be designed based on the nucleotide sequence information provided in any of SEQ ID NOs:1-169 (odd numbers). Typically, a primer consists of 10 to 40, preferably 15 to 25 nucleotides. It can also be advantageous to select primers containing C and G nucleotides in proportions sufficient to ensure efficient hybridization, *e.g.*, an amount of C and G nucleotides of at least 40%, preferably 50%, of the total nucleotide amount. Those skilled in the art can readily design primers that can be used to isolate the polynucleotides of the invention from different *Helicobacter* strains. Experimental conditions for carrying out PCR can readily be determined by one skilled in the art and an illustration of carrying out PCR is provided in the Examples below. As is well known in the art, restriction endonuclease recognition sites that contain, typically, 4 to 6 nucleotides (for example, the sequences 5'-GGATCC-3' (*Bam*HI) or 5'-CTCGAG-3' (*Xho*I)), can be included on the 5' ends of the primers. Restriction sites can be selected by those skilled in the art so that the amplified DNA can be conveniently cloned into an appropriately digested vector, such as a plasmid.

Useful homologs that do not occur naturally can be designed using known methods for identifying regions of an antigen that are likely to be tolerant of amino acid sequence changes and/or deletions. For example,

-14-

sequences of the antigen from different species can be compared to identify conserved sequences.

Polypeptide derivatives that are encoded by polynucleotides of the invention include, *e.g.*, fragments, polypeptides having large internal deletions
5 derived from full-length polypeptides, and fusion proteins. Polypeptide fragments of the invention can be derived from a polypeptide having a sequence homologous to any of the sequences of SEQ ID NOs:2-170 (even numbers), to the extent that the fragments retain the substantial antigenicity of the parent polypeptide (specific antigenicity). Polypeptide derivatives can also
10 be constructed by large internal deletions that remove a substantial part of the parent polypeptide, while retaining specific antigenicity. Generally, polypeptide derivatives should be about at least 12 amino acids in length to maintain antigenicity. Advantageously, they can be at least 20 amino acids, preferably at least 50 amino acids, more preferably at least 75 amino acids, and
15 most preferably at least 100 amino acids in length.

Useful polypeptide derivatives, *e.g.*, polypeptide fragments, can be designed using computer-assisted analysis of amino acid sequences in order to identify sites in protein antigens having potential as surface-exposed, antigenic regions (Hughes *et al.*, Infect. Immun. 60(9):3497, 1992). For example, the
20 Laser Gene Program from DNA Star can be used to obtain hydrophilicity, antigenic index, and intensity index plots for the polypeptides of the invention. This program can also be used to obtain information about homologies of the polypeptides with known protein motifs. One skilled in the art can readily use the information provided in such plots to select peptide fragments for use as
25 vaccine antigens. For example, fragments spanning regions of the plots in which the antigenic index is relatively high can be selected. One can also select fragments spanning regions in which both the antigenic index and the

-15-

intensity plots are relatively high. Fragments containing conserved sequences, particularly hydrophilic conserved sequences, can also be selected.

Polypeptide fragments and polypeptides having large internal deletions can be used for revealing epitopes that are otherwise masked in the parent polypeptide and that may be of importance for inducing a protective T cell-dependent immune response. Deletions can also remove immunodominant regions of high variability among strains.

It is an accepted practice in the field of immunology to use fragments and variants of protein immunogens as vaccines, as all that is required to induce an immune response to a protein is a small (*e.g.*, 8 to 10 amino acids) immunogenic region of the protein. This has been done for a number of vaccines against pathogens other than *Helicobacter*. For example, short synthetic peptides corresponding to surface-exposed antigens of pathogens such as murine mammary tumor virus (peptide containing 11 amino acids; Dion *et al.*, Virology 179:474-477, 1990), Semliki Forest virus (peptide containing 16 amino acids; Snijders *et al.*, J. Gen. Virol. 72:557-565, 1991), and canine parvovirus (2 overlapping peptides, each containing 15 amino acids; Langeveld *et al.*, Vaccine 12(15):1473-1480, 1994) have been shown to be effective vaccine antigens against their respective pathogens.

Polynucleotides encoding polypeptide fragments and polypeptides having large internal deletions can be constructed using standard methods (see, *e.g.*, Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons Inc., 1994), for example, by PCR, including inverse PCR, by restriction enzyme treatment of the cloned DNA molecules, or by the method of Kunkel *et al.* (Proc. Natl. Acad. Sci. USA 82:448, 1985; biological material available at Stratagene).

-16-

A polypeptide derivative can also be produced as a fusion polypeptide that contains a polypeptide or a polypeptide derivative of the invention fused, *e.g.*, at the N- or C-terminal end, to any other polypeptide (hereinafter referred to as a peptide tail). Such a product can be easily obtained
5 by translation of a genetic fusion, *i.e.*, a hybrid gene. Vectors for expressing fusion polypeptides are commercially available, and include the pMal-c2 or pMal-p2 systems of New England Biolabs, in which the peptide tail is a maltose binding protein, the glutathione-S-transferase system of Pharmacia, or the His-Tag system available from Novagen. These and other expression
10 systems provide convenient means for further purification of polypeptides and derivatives of the invention.

Another particular example of fusion polypeptides included in invention includes a polypeptide or polypeptide derivative of the invention fused to a polypeptide having adjuvant activity, such as, *e.g.*, subunit B of
15 either cholera toxin or *E. coli* heat-labile toxin. Several possibilities can be used for producing such fusion proteins. First, the polypeptide of the invention can be fused to the
N-terminal end or, preferably, to the C-terminal end of the polypeptide having adjuvant activity. Second, a polypeptide fragment of the invention can be fused
20 within the amino acid sequence of the polypeptide having adjuvant activity. Spacer sequences can also be included, if desired.

As stated above, the polynucleotides of the invention encode *Helicobacter* polypeptides in precursor or mature form. They can also encode hybrid precursors containing heterologous signal peptides, which can mature
25 into polypeptides of the invention. By "heterologous signal peptide" is meant a signal peptide that is not found in the naturally-occurring precursor of a polypeptide of the invention.

-17-

A polynucleotide of the invention hybridizes, preferably under stringent conditions, to a polynucleotide having a sequence as shown in any of SEQ ID NOs:1-169 (odd numbers). Hybridization procedures are, *e.g.*, described by Ausubel *et al.* (*supra*); Silhavy *et al.* (*Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1984); and Davis *et al.* (*A Manual for Genetic Engineering: Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1980). Important parameters that can be considered for optimizing hybridization conditions are reflected in the following formula, which facilitates calculation of the melting temperature (T_m), which is the temperature above which two complementary DNA strands separate from one another (Casey *et al.*, Nucl. Acid Res. 4:1539, 1977): $T_m = 81.5 + 0.5 \times (\% \text{ G+C}) + 1.6 \log (\text{positive ion concentration}) - 0.6 \times (\% \text{ formamide})$. Under appropriate stringency conditions, hybridization temperature (T_h) is approximately 20 to 40°C, 20 to 25°C, or, preferably, 30 to 40°C below the calculated T_m . Those skilled in the art will understand that optimal temperature and salt conditions can be readily determined empirically in preliminary experiments using conventional procedures. For example, stringent conditions can be achieved, both for pre-hybridizing and hybridizing incubations, (i) within 4-16 hours at 42°C, in 6 x SSC containing 50% formamide or (ii) within 4-16 hours at 65°C in an aqueous 6 x SSC solution (1 M NaCl, 0.1 M sodium citrate (pH 7.0)). For polynucleotides containing 30 to 600 nucleotides, the above formula is used and then is corrected by subtracting (600/polynucleotide size in base pairs). Stringency conditions are defined by a T_h that is 5 to 10°C below T_m .

Hybridization conditions with oligonucleotides shorter than 20-30 bases do not precisely follow the rules set forth above. In such cases, the

-18-

formula for calculating the T_m is as follows: $T_m = 4 \times (G+C) + 2 \times (A+T)$. For example, an 18 nucleotide fragment of 50% G+C would have an approximate T_m of 54°C.

A polynucleotide molecule of the invention, containing RNA, DNA,
5 or modifications or combinations thereof, can have various applications. For example, a polynucleotide molecule can be used (i) in a process for producing the encoded polypeptide in a recombinant host system, (ii) in the construction of vaccine vectors, such as poxviruses, which are further used in methods and compositions for preventing and/or treating *Helicobacter* infection, (iii) as a
10 vaccine agent, in a naked form or formulated with a delivery vehicle, and (iv) in the construction of attenuated *Helicobacter* strains that can over-express a polynucleotide of the invention or express it in a non-toxic, mutated form.

According to a second aspect of the invention, there is therefore provided (i) an expression cassette containing a polynucleotide molecule of the
15 invention placed under the control of elements (*e.g.*, a promoter) required for expression; (ii) an expression vector containing an expression cassette of the invention; (iii) a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, as well as (iv) a process for producing a polypeptide or polypeptide derivative encoded by a
20 polynucleotide of the invention, which involves culturing a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, under conditions that allow expression of the polynucleotide molecule of the invention, and recovering the encoded polypeptide or polypeptide derivative from the cell culture.

25 A recombinant expression system can be selected from procaryotic and eucaryotic hosts. Eucaryotic hosts include, for example, yeast cells (*e.g.*, *Saccharomyces cerevisiae* or *Pichia Pastoris*), mammalian cells (*e.g.*, COS1,

-19-

NIH3T3, or JEG3 cells), arthropods cells (*e.g.*, *Spodoptera frugiperda* (SF9) cells), and plant cells. Preferably, a procaryotic host such as *E. coli* is used. Bacterial and eucaryotic cells are available from a number of different sources that are known to those skilled in the art, *e.g.*, the American Type Culture
5 Collection (ATCC; Rockville, Maryland).

The choice of the expression cassette will depend on the host system selected, as well as the features desired for the expressed polypeptide. For example, it may be useful to produce a polypeptide of the invention in a particular lipidated form or any other form. Typically, an expression cassette
10 includes a constitutive or inducible promoter that is functional in the selected host system; a ribosome binding site; a start codon (ATG); if necessary, a region encoding a signal peptide, *e.g.*, a lipidation signal peptide; a polynucleotide molecule of the invention; a stop codon; and, optionally, a 3' terminal region (translation and/or transcription terminator). The signal
15 peptide-encoding region is adjacent to the polynucleotide of the invention and is placed in the proper reading frame. The signal peptide-encoding region can be homologous or heterologous to the polynucleotide molecule encoding the mature polypeptide and it can be specific to the secretion apparatus of the host used for expression. The open reading frame constituted by the polynucleotide
20 molecule of the invention, alone or together with the signal peptide, is placed under the control of the promoter so that transcription and translation occur in the host system. Promoters and signal peptide-encoding regions are widely known and available to those skilled in the art and include, for example, the promoter of *Salmonella typhimurium* (and derivatives) that is inducible by
25 arabinose (promoter *araB*) and is functional in Gram-negative bacteria such as *E. coli* (U.S. Patent No. 5,028,530; Cagnon *et al.*, Protein Engineering 4(7):843, 1991); the promoter of the bacteriophage T7 RNA polymerase gene,

-20-

which is functional in a number of *E. coli* strains expressing T7 polymerase (U.S. Patent No. 4,952,496); the OspA lipitation signal peptide; and RlpB lipitation signal peptide (Takase *et al.*, J. Bact. 169:5692, 1987).

The expression cassette is typically part of an expression vector,
5 which is selected for its ability to replicate in the chosen expression system. Expression vectors (*e.g.*, plasmids or viral vectors) can be chosen from, for example, those described in Pouwels *et al.* (*Cloning Vectors: A Laboratory Manual*, 1985, Supp. 1987), and can be purchased from various commercial sources. Methods for transforming or transfecting host cells with expression
10 vectors are well known in the art and will depend on the host system selected, as described in Ausubel *et al.* (*supra*).

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains in the intracellular compartment, is secreted/excreted in the extracellular medium or in the
15 periplasmic space, or is embedded in the cellular membrane. The polypeptide can then be recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the cell culture. Typically, the recombinant polypeptide can be purified by antibody-based affinity purification or by any other method known in the art, such as by genetic fusion to a small
20 affinity-binding domain. Antibody-based affinity purification methods are also available for purifying a polypeptide of the invention extracted from a *Helicobacter* strain. Antibodies useful for immunoaffinity purification of the polypeptides of the invention can be obtained using methods described below.

Polynucleotides of the invention can also be used in DNA
25 vaccination methods, using either a viral or bacterial host as gene delivery vehicle (live vaccine vector) or administering the gene in a free form, *e.g.*,

-21-

inserted into a plasmid. Therapeutic or prophylactic efficacy of a polynucleotide of the invention can be evaluated as is described below.

Accordingly, in a third aspect of the invention, there is provided (i) a vaccine vector such as a poxvirus, containing a polynucleotide molecule of the invention placed under the control of elements required for expression; (ii) a composition of matter containing a vaccine vector of the invention, together with a diluent or carrier; (iii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a vaccine vector of the invention; (iv) a method for inducing an immune response against *Helicobacter* in a mammal (*e.g.*, a human; alternatively, the method can be used in veterinary applications for treating or preventing *Helicobacter* infection of animals, *e.g.*, cats or birds), which involves administering to the mammal an immunogenically effective amount of a vaccine vector of the invention to elicit an immune response, *e.g.*, a protective or therapeutic immune response to *Helicobacter*; and (v) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, which involves administering a prophylactic or therapeutic amount of a vaccine vector of the invention to an individual in need. Additionally, the third aspect of the invention encompasses the use of a vaccine vector of the invention in the preparation of a medicament for preventing and/or treating *Helicobacter* infection.

A vaccine vector of the invention can express one or several polypeptides or derivatives of the invention, as well as at least one additional *Helicobacter* antigen such as a urease apoenzyme or a subunit, fragment, homolog, mutant, or derivative thereof. In addition, it can express a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), that enhances the immune response. Thus, a vaccine vector can include an additional

-22-

polynucleotide molecule encoding, *e.g.*, urease subunit A, B, or both, or a cytokine, placed under the control of elements required for expression in a mammalian cell.

Alternatively, a composition of the invention can include several
5 vaccine vectors, each of which are capable of expressing a polypeptide or derivative of the invention. A composition can also contain a vaccine vector capable of expressing an additional *Helicobacter* antigen, such as urease apoenzyme, a subunit, fragment, homolog, mutant, or derivative thereof, or a cytokine such as IL-2 or IL-12.

10 In vaccination methods for treating or preventing infection in a mammal, a vaccine vector of the invention can be administered by any conventional route in use in the vaccine field, for example, to a mucosal (*e.g.*, ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) surface or *via* a parenteral (*e.g.*, subcutaneous, intradermal,
15 intramuscular, intravenous, or intraperitoneal) route, or a combination thereof. Preferred routes depend upon the choice of the vaccine vector. The administration can be achieved in a single dose or repeated at intervals. The appropriate dosage depends on various parameters that are understood by those skilled in the art, such as the nature of the vaccine vector itself, the route of
20 administration, and the condition of the mammal to be vaccinated (*e.g.*, the weight, age, and general health of the mammal).

Live vaccine vectors that can be used in the invention include viral vectors, such as adenoviruses and poxviruses, as well as bacterial vectors, *e.g.*, *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-
25 Guérin (BCG), and *Streptococcus*. An example of an adenovirus vector, as well as a method for constructing an adenovirus vector capable of expressing a polynucleotide molecule of the invention, is described in U.S. Patent No.

-23-

4,920,209. Poxvirus vectors that can be used in the invention include, *e.g.*, vaccinia and canary pox viruses, which are described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively (also see, *e.g.*, Tartaglia *et al.*, Virology 188:217, 1992, for a description of a vaccinia virus vector, and
5 Taylor *et al.*, Vaccine 13:539, 1995, for a description of a canary poxvirus vector). Poxvirus vectors capable of expressing a polynucleotide of the invention can be obtained by homologous recombination, as described in Kieny *et al.* (Nature 312:163, 1984) so that the polynucleotide of the invention is inserted into the viral genome under appropriate conditions for expression in
10 mammalian cells. Generally, the dose of viral vector vaccine, for therapeutic or prophylactic use, can be from about 1×10^4 to about 1×10^{11} , advantageously from about 1×10^7 to about 1×10^{10} , or, preferably, from about 1×10^7 to about 1×10^9 plaque-forming units per kilogram. Preferably, viral vectors are administered parenterally, for example, in 3 doses that are 4 weeks apart.
15 Those skilled in the art will recognize that it is preferable to avoid adding a chemical adjuvant to a composition containing a viral vector of the invention and thereby minimizing the immune response to the viral vector itself.

Non-toxicogenic *Vibrio cholerae* mutant strains that can be used in live oral vaccines are described by Mekalanos *et al.* (Nature 306:551, 1983)
20 and in U.S. Patent No. 4,882,278 (strain in which a substantial amount of the coding sequence of each of the two *ctxA* alleles has been deleted so that no functional *cholerae* toxin is produced); WO 92/11354 (strain in which the *irgA* locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA* mutations); and WO 94/1533 (deletion mutant lacking
25 functional *ctxA* and *attRSI* DNA sequences). These strains can be genetically engineered to express heterologous antigens, as described in WO 94/19482. An effective vaccine dose of a *V. cholerae* strain capable of expressing a

-24-

polypeptide or polypeptide derivative encoded by a polynucleotide molecule of the invention can contain, *e.g.*, about 1×10^5 to about 1×10^9 , preferably about 1×10^6 to about 1×10^8 , viable bacteria in an appropriate volume for the selected route of administration. Preferred routes of administration include all mucosal routes, but, most preferably, these vectors are administered intranasally or orally.

Attenuated *Salmonella typhimurium* strains, genetically engineered for recombinant expression of heterologous antigens, and their use as oral vaccines, are described by Nakayama *et al.* (Bio/Technology 6:693, 1988) and in WO 92/11361. Preferred routes of administration for these vectors include all mucosal routes. Most preferably, the vectors are administered intranasally or orally.

Others bacterial strains useful as vaccine vectors are described by High *et al.* (EMBO 11:1991, 1992) and Sizemore *et al.* (Science 270:299, 1995; *Shigella flexneri*); Medaglini *et al.* (Proc. Natl. Acad. Sci. USA 92:6868, 1995; (*Streptococcus gordonii*); Flynn (Cell. Mol. Biol. 40 (suppl. I):31, 1194), and in WO 88/6626, WO 90/0594, WO 91/13157, WO 92/1796, and WO 92/21376 (Bacille Calmette Guerin). In bacterial vectors, a polynucleotide of the invention can be inserted into the bacterial genome or it can remain in a free state, for example, carried on a plasmid.

An adjuvant can also be added to a composition containing a bacterial vector vaccine. A number of adjuvants that can be used are known to those skilled in the art. For example, preferred adjuvants can be selected from the list provided below.

According to a fourth aspect of the invention, there is also provided (i) a composition of matter containing a polynucleotide of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition containing

-25-

a therapeutically or prophylactically effective amount of a polynucleotide of the invention; (iii) a method for inducing an immune response against *Helicobacter* in a mammal by administering to the mammal an immunogenically effective amount of a polynucleotide of the invention to elicit an immune response, *e.g.*,
5 a protective immune response to *Helicobacter*; and (iv) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a prophylactic or therapeutic amount of a polynucleotide of the invention to an individual in need of such treatment. Additionally, the fourth aspect of the invention encompasses the use
10 of a polynucleotide of the invention in the preparation of a medicament for preventing and/or treating *Helicobacter* infection. The fourth aspect of the invention preferably includes the use of a polynucleotide molecule placed under conditions for expression in a mammalian cell, *e.g.*, in a plasmid that is unable to replicate in mammalian cells and to substantially integrate into a
15 mammalian genome.

Polynucleotides (for example, DNA or RNA molecules) of the invention can also be administered as such to a mammal as a vaccine. When a DNA molecule of the invention is used, it can be in the form of a plasmid that is unable to replicate in a mammalian cell and unable to integrate into the
20 mammalian genome. Typically, a DNA molecule is placed under the control of a promoter suitable for expression in a mammalian cell. The promoter can function ubiquitously or tissue-specifically. Examples of non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (U.S. Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (Norton *et al.*, Molec.
25 Cell Biol. 5:281, 1985). The desmin promoter (Li *et al.*, Gene 78:243, 1989; Li *et al.*, J. Biol. Chem. 266:6562, 1991; Li *et al.*, J. Biol. Chem. 268:10403, 1993) is tissue-specific and drives expression in muscle cells. More generally,

-26-

useful promoters and vectors are described, *e.g.*, in WO 94/21797 and by Hartikka *et al.* (Human Gene Therapy 7:1205, 1996).

For DNA/RNA vaccination, the polynucleotide of the invention can encode a precursor or a mature form of a polypeptide of the invention. When it
5 encodes a precursor form, the precursor sequence can be homologous or heterologous. In the latter case, a eucaryotic leader sequence can be used, such as the leader sequence of the tissue-type plasminogen factor (tPA).

A composition of the invention can contain one or several polynucleotides of the invention. It can also contain at least one additional
10 polynucleotide encoding another *Helicobacter* antigen, such as urease subunit A, B, or both, or a fragment, derivative, mutant, or analog thereof. A polynucleotide encoding a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), can also be added to the composition so that the immune response is enhanced. These additional polynucleotides are placed under appropriate
15 control for expression. Advantageously, DNA molecules of the invention and/or additional DNA molecules to be included in the same composition are carried in the same plasmid.

Standard methods can be used in the preparation of therapeutic polynucleotides of the invention. For example, a polynucleotide can be used in
20 a naked form, free of any delivery vehicles, such as anionic liposomes, cationic lipids, microparticles, *e.g.*, gold microparticles, precipitating agents, *e.g.*, calcium phosphate, or any other transfection-facilitating agent. In this case, the polynucleotide can be simply diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline, with or without a carrier. When
25 present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, *e.g.*, a solution containing 20% sucrose.

-27-

Alternatively, a polynucleotide can be associated with agents that assist in cellular uptake. It can be, *e.g.*, (i) complemented with a chemical agent that modifies cellular permeability, such as bupivacaine (see, *e.g.*, WO 94/16737), (ii) encapsulated into liposomes, or (iii) associated with cationic lipids or silica, gold, or tungsten microparticles.

Anionic and neutral liposomes are well-known in the art (see, *e.g.*, *Liposomes: A Practical Approach*, RPC New Ed, IRL Press, 1990, for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides.

Cationic lipids can also be used for gene delivery. Such lipids include, for example, Lipofectin™, which is also known as DOTMA (N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-bis(oleyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidoglycyl spermine), and cholesterol derivatives. A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids for gene delivery are preferably used in association with a neutral lipid, such as DOPE (dioleoyl phosphatidylethanolamine; WO 90/11092). Other transfection-facilitating compounds can be added to a formulation containing cationic liposomes. A number of them are described in, *e.g.*, WO 93/18759, WO 93/19768, WO 94/25608, and WO 95/2397. They include, *e.g.*, spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example, WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles can also be used for gene delivery, as described in WO 91/359, WO 93/17706, and by Tang *et al.* (Nature 356:152, 1992). In this case, the microparticle-coated polynucleotides can be injected *via* intradermal or intraepidermal routes using a needleless injection device

5 (“gene gun”), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

The amount of DNA to be used in a vaccine depends, *e.g.*, on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for

10 administration (*e.g.*, the weight, age, and general health of the mammal), the mode of administration, and the type of formulation. In general, a therapeutically or prophylactically effective dose from about 1 μ g to about 1 mg, preferably, from about 10 μ g to about 800 μ g, and, more preferably, from about 25 μ g to about 250 μ g, can be administered to a human adult. The

15 administration can be achieved in a single dose or repeated at intervals.

The route of administration can be any conventional route used in the vaccine field. As general guidance, a polynucleotide of the invention can be administered *via* a mucosal surface, *e.g.*, an ocular, intranasal, pulmonary, oral, intestinal, rectal, vaginal, or urinary tract surface, or *via* a parenteral route, *e.g.*,

20 by an intravenous, subcutaneous, intraperitoneal, intradermal, intraepidermal, or intramuscular route. The choice of administration route will depend on, *e.g.*, the formulation that is selected. A polynucleotide formulated in association with bupivacaine is advantageously administered into muscle. When a neutral or anionic liposome or a cationic lipid, such as DOTMA, is used, the

25 formulation can be advantageously administered *via* intravenous, intranasal (for example, by aerosolization), intramuscular, intradermal, and subcutaneous routes. A polynucleotide in a naked form can advantageously be administered

-29-

via the intramuscular, intradermal, or subcutaneous routes. Although not absolutely required, such a composition can also contain an adjuvant. A systemic adjuvant that does not require concomitant administration in order to exhibit an adjuvant effect is preferable.

5 The sequence information provided in the present application enables the design of specific nucleotide probes and primers that can be used in diagnostic methods. Accordingly, in a fifth aspect of the invention, there is provided a nucleotide probe or primer having a sequence found in, or derived by degeneracy of the genetic code from, a sequence shown in any of SEQ ID
10 NOs:1-169 (odd numbers).

 The term "probe" as used in the present application refers to a DNA (preferably single stranded) or RNA molecule (or modifications or combinations thereof) that hybridizes under the stringent conditions, as defined above, to a polynucleotide molecule having a sequence homologous to any of
15 those shown in SEQ ID NOs:1-169 (odd numbers), or to a complementary or anti-sense sequence of any of those shown in SEQ ID NOs:1-169 (odd numbers). Generally, probes are significantly shorter than the full-length sequences shown in SEQ ID NOs:1-169 (odd numbers). For example, they can contain from about 5 to about 100, preferably from about 10 to about 80,
20 nucleotides. In particular, probes have sequences that are at least 75%, preferably at least 85%, more preferably 95% homologous to a portion of a sequence as shown in any of SEQ ID NOs:1-169 (odd numbers) or a sequence complementary to any of such sequences.

 Probes can contain modified bases, such as inosine, methyl-5-
25 deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine, or diamino-2, 6-purine. Sugar or phosphate residues can also be modified or substituted. For example, a deoxyribose residue can be replaced by a polyamide (Nielsen *et al.*,

-30-

Science 254:1497, 1991) and phosphate residues can be replaced by ester groups, such as diphosphate, alkyl, arylphosphonate, and phosphorothioate esters. In addition, the 2'-hydroxyl group on ribonucleotides can be modified by addition of, *e.g.*, alkyl groups.

5 Probes of the invention can be used in diagnostic tests or as capture or detection probes. Such capture probes can be immobilized on solid supports, directly or indirectly, by covalent means or by passive adsorption. A detection probe can be labeled by a detectable label, for example, a label selected from radioactive isotopes; enzymes, such as peroxidase and alkaline phosphatase;
10 enzymes that are able to hydrolyze a chromogenic, fluorogenic, or luminescent substrate; compounds that are chromogenic, fluorogenic, or luminescent; nucleotide base analogs; and biotin.

 Probes of the invention can be used in any conventional hybridization method, such as in dot blot methods (Maniatis *et al.*, *Molecular*
15 *Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1982), Southern blot methods (Southern, J. Mol. Biol. 98:503, 1975), northern blot methods (identical to Southern blot to the exception that RNA is used as a target), or a sandwich method (Dunn *et al.*, Cell 12:23, 1977). As is known in the art, the latter technique involves the use
20 of a specific capture probe and a specific detection probe that have nucleotide sequences that are at least partially different from each other.

 Primers used in the invention usually contain about 10 to 40 nucleotides and are used to initiate enzymatic polymerization of DNA in an amplification process (*e.g.*, PCR), an elongation process, or a reverse
25 transcription method. In a diagnostic method involving PCR, the primers can be labeled.

-31-

Thus, the invention also encompasses (i) a reagent containing a probe of the invention for detecting and/or identifying the presence of *Helicobacter* in a biological material; (ii) a method for detecting and/or identifying the presence of *Helicobacter* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA or RNA is extracted from the material and denatured, and (c) the sample is exposed to a probe of the invention, for example, a capture probe, a detection probe, or both, under stringent hybridization conditions, so that hybridization is detected; and (iii) a method for detecting and/or identifying the presence of *Helicobacter* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA is extracted therefrom, (c) the extracted DNA is contacted with at least one, or, preferably two, primers of the invention, and amplified by the polymerase chain reaction, and (d) an amplified DNA molecule is produced.

As mentioned above, polypeptides that can be produced by expression of the polynucleotides of the invention can be used as vaccine antigens. Accordingly, a sixth aspect of the invention features a substantially purified polypeptide or polypeptide derivative having an amino acid sequence encoded by a polynucleotide of the invention.

A "substantially purified polypeptide" is defined as a polypeptide that is separated from the environment in which it naturally occurs and/or a polypeptide that is free of most of the other polypeptides that are present in the environment in which it was synthesized. The polypeptides of the invention can be purified from a natural source, such as a *Helicobacter* strain, or can be produced using recombinant methods.

Homologous polypeptides or polypeptide derivatives encoded by polynucleotides of the invention can be screened for specific antigenicity by

-32-

testing cross-reactivity with an antiserum raised against a polypeptide having an amino acid sequence as shown in any of SEQ ID NOs:2-170 (even numbers). Briefly, a monospecific hyperimmune antiserum can be raised against a purified reference polypeptide as such or as a fusion polypeptide, for example, an expression product of MBP, GST, or His-tag systems, or a synthetic peptide predicted to be antigenic. The homologous polypeptide or derivative that is screened for specific antigenicity can be produced as such or as a fusion polypeptide. In the latter case, and if the antiserum is also raised against a fusion polypeptide, two different fusion systems are employed.

Specific antigenicity can be determined using a number of methods, including Western blot (Towbin *et al.*, Proc. Natl. Acad. Sci. USA 76:4350, 1979), dot blot, and ELISA methods, as described below.

In a Western blot assay, the product to be screened, either as a purified preparation or a total *E. coli* extract, is fractionated by SDS-PAGE, as described, for example, by Laemmli (Nature 227:680, 1970). After being transferred to a filter, such as a nitrocellulose membrane, the material is incubated with the monospecific hyperimmune antiserum, which is diluted in a range of dilutions from about 1:50 to about 1:5,000, preferably from about 1:100 to about 1:500. Specific antigenicity is shown once a band corresponding to the product exhibits reactivity at any of the dilutions in the range.

In an ELISA assay, the product to be screened can be used as the coating antigen. A purified preparation is preferred, but a whole cell extract can also be used. Briefly, about 100 μ l of a preparation of about 10 μ g protein/ml is distributed into wells of a 96-well ELISA plate. The plate is incubated for about 2 hours at 37°C, then overnight at 4°C. The plate is washed with phosphate buffer saline (PBS) containing 0.05% Tween 20

-33-

(PBS/Tween buffer) and the wells are saturated with 250 μ l PBS containing 1% bovine serum albumin (BSA), to prevent non-specific antibody binding. After 1 hour of incubation at 37°C, the plate is washed with PBS/Tween buffer. The antiserum is serially diluted in PBS/Tween buffer containing 0.5% BSA, and 100 μ l dilutions are added to each well. The plate is incubated for 90 minutes at 37°C, washed, and evaluated using standard methods. For example, a goat anti-rabbit peroxidase conjugate can be added to the wells when the specific antibodies used were raised in rabbits. Incubation is carried out for about 90 minutes at 37°C and the plate is washed. The reaction is developed with the appropriate substrate and the reaction is measured by colorimetry (absorbance measured spectrophotometrically). Under these experimental conditions, a positive reaction is shown once an O.D. value of 1.0 is detected with a dilution of at least about 1:50, preferably of at least about 1:500.

In a dot blot assay, a purified product is preferred, although a whole cell extract can be used. Briefly, a solution of the product at a concentration of about 100 μ g/ml is serially diluted two-fold with 50 mM Tris-HCl (pH 7.5). One hundred μ l of each dilution is applied to a filter, such as a 0.45 μ m nitrocellulose membrane, set in a 96-well dot blot apparatus (Biorad). The buffer is removed by applying vacuum to the system. Wells are washed by addition of 50 mM Tris-HCl (pH 7.5) and the membrane is air-dried. The membrane is saturated in blocking buffer (50 mM Tris-HCl (pH 7.5), 0.15 M NaCl, 10 g/l skim milk) and incubated with an antiserum diluted from about 1:50 to about 1:5000, preferably about 1:500. The reaction is detected using standard methods. For example, a goat anti-rabbit peroxidase conjugate can be added to the wells when rabbit antibodies are used. Incubation is carried out for about 90 minutes at 37°C and the blot is washed. The reaction is developed

-34-

with the appropriate substrate and stopped. The reaction is then measured visually by the appearance of a colored spot, *e.g.*, by colorimetry. Under these experimental conditions, a positive reaction is associated with detection of a colored spot for reactions carried out with a dilution of at least about 1:50, preferably, of at least about 1:500. Therapeutic or prophylactic efficacy of a polypeptide or polypeptide derivative of the invention can be evaluated as described below.

According to a seventh aspect of the invention, there is provided (i) a composition of matter containing a polypeptide of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polypeptide of the invention; (iii) a method for inducing an immune response against *Helicobacter* in a mammal by administering to the mammal an immunogenically effective amount of a polypeptide of the invention to elicit an immune response, *e.g.*, a protective immune response to *Helicobacter*; and (iv) a method for preventing and/or treating a *Helicobacter* (*e.g.*, *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a prophylactic or therapeutic amount of a polypeptide of the invention to an individual in need of such treatment. Additionally, this aspect of the invention includes the use of a polypeptide of the invention in the preparation of a medicament for preventing and/or treating *Helicobacter* infection.

The immunogenic compositions of the invention can be administered by any conventional route in use in the vaccine field, for example, to a mucosal (*e.g.*, ocular, intranasal, pulmonary, oral, gastric, intestinal, rectal, vaginal, or urinary tract) surface or *via* a parenteral (*e.g.*, subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. The choice of the administration route depends upon a number of parameters, such as the

-35-

adjuvant used. For example, if a mucosal adjuvant is used, the intranasal or oral route will be preferred, and if a lipid formulation or an aluminum compound is used, a parenteral route will be preferred. In the latter case, the subcutaneous or intramuscular route is most preferred. The choice of administration route can also depend upon the nature of the vaccine agent. For example, a polypeptide of the invention fused to CTB or to LTB will be best administered to a mucosal surface.

A composition of the invention can contain one or several polypeptides or derivatives of the invention. It can also contain at least one additional *Helicobacter* antigen, such as the urease apoenzyme, or a subunit, fragment, homolog, mutant, or derivative thereof.

For use in a composition of the invention, a polypeptide or polypeptide derivative can be formulated into or with liposomes, such as neutral or anionic liposomes, microspheres, ISCOMS, or virus-like particles (VLPs), to facilitate delivery and/or enhance the immune response. These compounds are readily available to those skilled in the art; for example, see *Liposomes: A Practical Approach (supra)*. Adjuvants other than liposomes can also be used in the invention and are well known in the art (see, for example, the list provided below).

Administration can be achieved in a single dose or repeated as necessary at appropriate intervals that can be determined by those skilled in the art. For example, a priming dose can be followed by three booster doses at weekly or monthly intervals. An appropriate dose depends on various parameters, including the nature of the recipient (*e.g.*, whether the recipient is an adult or an infant), the particular vaccine antigen, the route and frequency of administration, the presence/absence or type of adjuvant, and the desired effect (*e.g.*, protection and/or treatment), and can be readily determined by one skilled

-36-

in the art. In general, a vaccine antigen of the invention can be administered mucosally in an amount ranging from about 10 μ g to about 500 mg, preferably from about 1 mg to about 200 mg. For a parenteral route of administration, the dose usually should not exceed about 1 mg, and is, preferably, about 100 μ g.

5 When used as components of a vaccine, the polynucleotides and polypeptides of the invention can be used sequentially as part of a multi-step immunization process. For example, a mammal can be initially primed with a vaccine vector of the invention, such as a pox virus, *e.g.*, *via* a parenteral route, and then boosted twice with a polypeptide encoded by the vaccine vector, *e.g.*,
10 *via* the mucosal route. In another example, liposomes associated with a polypeptide or polypeptide derivative of the invention can be used for priming, with boosting being carried out mucosally using a soluble polypeptide or polypeptide derivative of the invention, in combination with a mucosal adjuvant (*e.g.*, LT).

15 Polypeptides and polypeptide derivatives of the invention can also be used as diagnostic reagents for detecting the presence of anti-*Helicobacter* antibodies, *e.g.*, in blood samples. Such polypeptides can be about 5 to about 80, preferably, about 10 to about 50, amino acids in length and can be labeled or unlabeled, depending upon the diagnostic method. Diagnostic methods
20 involving such a reagent are described below.

 Upon expression of a polynucleotide molecule of the invention, a polypeptide or polypeptide derivative is produced and can be purified using known methods. For example, the polypeptide or polypeptide derivative can be produced as a fusion protein containing a fused tail that facilitates purification.
25 The fusion product can be used to immunize a small mammal, *e.g.*, a mouse or a rabbit, in order to raise monospecific antibodies against the polypeptide or polypeptide derivative. The eighth aspect of the invention thus provides a

-37-

monospecific antibody that binds to a polypeptide or polypeptide derivative of the invention.

By "monospecific antibody" is meant an antibody that is capable of reacting with a unique, naturally-occurring *Helicobacter* polypeptide. An antibody of the invention can be polyclonal or monoclonal. Monospecific antibodies can be recombinant, *e.g.*, chimeric (*e.g.*, consisting of a variable region of murine origin and a human constant region), humanized (*e.g.*, a human immunoglobulin constant region and a variable region of animal, *e.g.*, murine, origin), and/or single chain. Both polyclonal and monospecific antibodies can also be in the form of immunoglobulin fragments, *e.g.*, F(ab)'2 or Fab fragments. The antibodies of the invention can be of any isotype, *e.g.*, IgG or IgA, and polyclonal antibodies can be of a single isotype or can contain a mixture of isotypes.

The antibodies of the invention, which can be raised against a polypeptide or polypeptide derivative of the invention, can be produced and identified using standard immunological assays, *e.g.*, Western blot assays, dot blot assays, or ELISA (see, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology*, John Wiley & Sons, Inc., New York, NY, 1994). The antibodies can be used in diagnostic methods to detect the presence of *Helicobacter* antigens in a sample, such as a biological sample. The antibodies can also be used in affinity chromatography methods for purifying a polypeptide or polypeptide derivative of the invention. As is discussed further below, the antibodies can also be used in prophylactic and therapeutic passive immunization methods.

Accordingly, a ninth aspect of the invention provides (i) a reagent for detecting the presence of *Helicobacter* in a biological sample that contains an antibody, polypeptide, or polypeptide derivative of the invention; and (ii) a

-38-

diagnostic method for detecting the presence of *Helicobacter* in a biological sample, by contacting the biological sample with an antibody, a polypeptide, or a polypeptide derivative of the invention, so that an immune complex is formed, and detecting the complex as an indication of the presence of

5 *Helicobacter* in the sample or the organism from which the sample was derived. The immune complex is formed between a component of the sample and the antibody, polypeptide, or polypeptide derivative, and any unbound material can be removed prior to detecting the complex. A polypeptide reagent can be used for detecting the presence of anti-*Helicobacter* antibodies in a
10 sample, *e.g.*, a blood sample, while an antibody of the invention can be used for screening a sample, such as a gastric extract or biopsy sample, for the presence of *Helicobacter* polypeptides.

For use in diagnostic methods, the reagent (*e.g.*, the antibody, polypeptide, or polypeptide derivative of the invention) can be in a free state or
15 can be immobilized on a solid support, such as, for example, on the interior surface of a tube or on the surface, or within pores, of a bead. Immobilization can be achieved using direct or indirect means. Direct means include passive adsorption (*i.e.*, non-covalent binding) or covalent binding between the support and the reagent. By "indirect means" is meant that an anti-reagent compound
20 that interacts with the reagent is first attached to the solid support. For example, if a polypeptide reagent is used, an antibody that binds to it can serve as an anti-reagent, provided that it binds to an epitope that is not involved in recognition of antibodies in biological samples. Indirect means can also employ a ligand-receptor system, for example, a molecule, such as a vitamin,
25 can be grafted onto the polypeptide reagent and the corresponding receptor can be immobilized on the solid phase. This concept is illustrated by the well known biotin-streptavidin system. Alternatively, indirect means can be used,

-39-

e.g., by adding to the reagent a peptide tail, chemically or by genetic engineering, and immobilizing the grafted or fused product by passive adsorption or covalent linkage of the peptide tail.

According to a tenth aspect of the invention, there is provided a
5 process for purifying from a biological sample a polypeptide or polypeptide derivative of the invention, which involves carrying out antibody-based affinity chromatography with the biological sample, wherein the antibody is a monospecific antibody of the invention.

For use in a purification process of the invention, the antibody can be
10 polyclonal or monospecific, and preferably is of the IgG type. Purified IgGs can be prepared from an antiserum using standard methods (see, *e.g.*, Coligan *et al.*, *supra*). Conventional chromatography supports, as well as standard methods for grafting antibodies, are described, for example, by Harlow *et al.* (*Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold
15 Spring Harbor, New York, 1988).

Briefly, a biological sample, such as an *H. pylori* extract, preferably in a buffer solution, is applied to a chromatography material, which is, preferably, equilibrated with the buffer used to dilute the biological sample, so that the polypeptide or polypeptide derivative of the invention (*i.e.*, the antigen)
20 is allowed to adsorb onto the material. The chromatography material, such as a gel or a resin coupled to an antibody of the invention, can be in batch form or in a column. The unbound components are washed off and the antigen is eluted with an appropriate elution buffer, such as a glycine buffer, a buffer containing a chaotropic agent, *e.g.*, guanidine HCl, or a buffer having high salt
25 concentration (*e.g.*, 3 M MgCl₂). Eluted fractions are recovered and the presence of the antigen is detected, *e.g.*, by measuring the absorbance at 280 nm.

-40-

An antibody of the invention can be screened for therapeutic efficacy as follows. According to an eleventh aspect of the invention, there is provided (i) a composition of matter containing a monospecific antibody of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition
5 containing a therapeutically or prophylactically effective amount of a monospecific antibody of the invention, and (iii) a method for treating or preventing *Helicobacter* (e.g., *H. pylori*, *H. felis*, *H. mustelae*, or *H. heilmanii*) infection, by administering a therapeutic or prophylactic amount of a monospecific antibody of the invention to an individual in need of such
10 treatment. In addition, the eleventh aspect of the invention includes the use of a monospecific antibody of the invention in the preparation of a medicament for treating or preventing *Helicobacter* infection.

The monospecific antibody can be polyclonal or monoclonal, and is, preferably, predominantly of the IgA isotype. In passive immunization
15 methods, the antibody is administered to a mucosal surface of a mammal, e.g., the gastric mucosa, e.g., orally or intragastrically, optionally, in the presence of a bicarbonate buffer. Alternatively, systemic administration, not requiring a bicarbonate buffer, can be carried out. A monospecific antibody of the invention can be administered as a single active agent or as a mixture with at
20 least one additional monospecific antibody specific for a different *Helicobacter* polypeptide. The amount of antibody and the particular regimen used can be readily determined by one skilled in the art. For example, daily administration of about 100 to 1,000 mg of antibody over one week, or three doses per day of about 100 to 1,000 mg of antibody over two or three days, can be effective
25 regimens for most purposes.

Therapeutic or prophylactic efficacy can be evaluated using standard methods in the art, e.g., by measuring induction of a mucosal immune response

-41-

or induction of protective and/or therapeutic immunity using, *e.g.*, the *H. felis* mouse model and the procedures described by Lee *et al.* (Eur. J. Gastroenterology & Hepatology 7:303, 1995) or Lee *et al.* (J. Infect. Dis. 172:161, 1995). Those skilled in the art will recognize that the *H. felis* strain of
5 the model can be replaced with another *Helicobacter* strain. For example, the efficacy of polynucleotide molecules and polypeptides from *H. pylori* is, preferably, evaluated in a mouse model using an *H. pylori* strain. Protection can be determined by comparing the degree of *Helicobacter* infection in the gastric tissue assessed by, for example, urease activity, bacterial counts, or
10 gastritis, to that of a control group. Protection is shown when infection is reduced by comparison to the control group. Such an evaluation can be made for polynucleotides, vaccine vectors, polypeptides, and polypeptide derivatives, as well as for antibodies of the invention.

For example, various doses of an antibody of the invention can be
15 administered to the gastric mucosa of mice previously challenged with an *H. pylori* strain as described, *e.g.*, by Lee *et al.* (*supra*). Then, after an appropriate period of time, the bacterial load of the mucosa can be estimated by assessing urease activity, as compared to a control. Reduced urease activity indicates that the antibody is therapeutically effective.

20 Adjuvants that can be used in any of the vaccine compositions described above are described as follows. Adjuvants for parenteral administration include, for example, aluminum compounds, such as aluminum hydroxide, aluminum phosphate, and aluminum hydroxy phosphate. The antigen can be precipitated with, or adsorbed onto, the aluminum compound
25 using standard methods. Other adjuvants, such as RIBI (ImmunoChem, Hamilton, MT), can also be used in parenteral administration.

-42-

Adjuvants that can be used for mucosal administration include, for example, bacterial toxins, *e.g.*, the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A, the *pertussis* toxin (PT), and combinations, subunits, toxoids, or mutants thereof. For example, a purified
5 preparation of native cholera toxin subunit B (CTB) can be used. Fragments, homologs, derivatives, and fusions to any of these toxins can also be used, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is used. Suitable mutants are described, *e.g.*, in WO 95/17211 (Arg-7-Lys CT mutant), WO 96/6627 (Arg-192-Gly LT mutant), and WO
10 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that can be used in the methods and compositions of the invention include, *e.g.*, Ser-63-Lys, Ala-69-Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants, such as the bacterial monophosphoryl lipid A (MPLA) of, *e.g.*, *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*;
15 saponins, and polylactide glycolide (PLGA) microspheres, can also be used in mucosal administration. Adjuvants useful for both mucosal and parenteral administration, such as polyphosphazene (WO 95/2415), can also be used.

Any pharmaceutical composition of the invention, containing a polynucleotide, polypeptide, polypeptide derivative, or antibody of the
20 invention, can be manufactured using standard methods. It can be formulated with a pharmaceutically acceptable diluent or carrier, *e.g.*, water or a saline solution, such as phosphate buffered saline, optionally, including a bicarbonate salt, such as sodium bicarbonate, *e.g.*, 0.1 to 0.5 M. Bicarbonate can advantageously be added to compositions intended for oral or intragastric
25 administration. In general, a diluent or carrier can be selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable-pharmaceutical carriers and diluents, as well as pharmaceutical

-43-

necessities for their use in pharmaceutical formulations, are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field and in the USP/NF.

The invention also includes methods in which gastroduodenal

5 infections, such as *Helicobacter* infection, are treated by oral administration of a *Helicobacter* polypeptide of the invention and a mucosal adjuvant, in combination with an antibiotic, an antisecretory agent, a bismuth salt, an antacid, sucralfate, or a combination thereof. Examples of such compounds that can be administered with the vaccine antigen and an adjuvant are

10 antibiotics, including, *e.g.*, macrolides, tetracyclines, β -lactams, aminoglycosides, quinolones, penicillins, and derivatives thereof (specific examples of antibiotics that can be used in the invention include, *e.g.*, amoxicillin, clarithromycin, tetracycline, metronidazole, erythromycin, cefuroxime, and erythromycin); antisecretory agents, including, *e.g.*, H₂-

15 receptor antagonists (*e.g.*, cimetidine, ranitidine, famotidine, nizatidine, and roxatidine), proton pump inhibitors (*e.g.*, omeprazole, lansoprazole, and pantoprazole), prostaglandin analogs (*e.g.*, misoprostil and enprostil), and anticholinergic agents (*e.g.*, pirenzepine, telenzepine, carbenoxolone, and proglumide); and bismuth salts, including colloidal bismuth subcitrate,

20 tripotassium dicitrate bismuthate, bismuth subsalicylate, bicitropeptide, and pepto-bismol (see, *e.g.*, Goodwin *et al.*, *Helicobacter pylori*, *Biology and Clinical Practice*, CRC Press, Boca Raton, FL, pp 366-395, 1993; Physicians' Desk Reference, 49th edn., Medical Economics Data Production Company, Montvale, New Jersey, 1995). In addition, compounds containing more than

25 one of the above-listed components coupled together, *e.g.*, ranitidine coupled to bismuth subcitrate, can be used. The invention also includes compositions for carrying out these methods, *i.e.*, compositions containing a *Helicobacter*

-44-

antigen (or antigens) of the invention, an adjuvant, and one or more of the above-listed compounds, in a pharmaceutically acceptable carrier or diluent.

Amounts of the above-listed compounds used in the methods and compositions of the invention can readily be determined by one skilled in the art. In addition, one skilled in the art can readily design treatment/immunization schedules. For example, the non-vaccine components can be administered on days 1-14, and the vaccine antigen + adjuvant can be administered on days 7, 14, 21, and 28.

Methods and pharmaceutical compositions of the invention can be used to treat or to prevent *Helicobacter* infections and, accordingly, gastroduodenal diseases associated with these infections, including acute, chronic, and atrophic gastritis, and peptic ulcer diseases, e.g., gastric and duodenal ulcers.

The clones of the invention were originally isolated by a transposon shuttle mutagenesis method. Briefly, in this method, a TnMax9 mini-*blaM* transposon was used for insertional mutagenesis of an *H. pylori* gene library established in *E. coli*. 192 *E. coli* clones expressing active β -lactamase fusion proteins were obtained, indicating that the corresponding target plasmids carry *H. pylori* genes encoding extracytoplasmic proteins. Individual mutants were transferred onto the chromosome of *H. pylori* P1 or P12 by natural transformation, resulting in 135 distinct *H. pylori* mutants. This method is described in further detail, as follows.

The transposon TnMax9 (Kahrs *et al.*, Gene 167:53, 1995) was used to generate mutations in an *H. pylori* library in *E. coli*. As illustrated in Fig. 1A, TnMax9 contains, in addition to a cat_{GC} -resistance gene close to the inverted repeat (IR), an unexpressed open reading frame encoding β -lactamase without a promoter or signal sequence (mature β -lactamase, *blaM*; Kahrs *et al.*,

-45-

supra). For production of extracytoplasmic BlaM fusion proteins resulting in ampicillin-resistant (amp^R) clones, expression of the cloned *H. pylori* genes in *E. coli* is obligatory. The minimal vector pMin2 (Kahrs *et al.*, *supra*; see Fig. 1B), containing a weak constitutive promoter (P_{iga}) upstream of the multiple cloning site, was used for construction of the *H. pylori* library to ensure expression of *H. pylori* genes in *E. coli*.

In construction of the library, *H. pylori* DNA was partially digested with *Sau3A* and *HpaII*, size fractionated by preparative agarose gel electrophoresis, and 3-6 kilobase fragments were ligated into the *BglII* and *ClaI* sites of pMin2. The library was introduced into *E. coli* strain E181(pTnMax9), which is a derivative of HB101 containing the TnMax9 transposon, by electroporation. This generated approximately 2,400 independent transformants. More than 95% of the plasmids contained an insert of between 3 and 6 kilobases, showing that the 1.7 megabase *H. pylori* chromosome was statistically covered. Since not every plasmid could be expected to contain a target gene carrying an export signal, the library was partitioned into a total of 198 pools (24 pools of 20 clones and 174 pools of 11 clones). Using a cotton swab, either eleven or twenty individual colonies were inoculated in 0.5 ml LB medium in eppendorf tubes, vortexed, and 100 μl of the suspension was spread on LB agar plates supplemented with tetracycline and chloramphenicol to select for maintenance of both plasmids. Insertion of TnMax9 into the target plasmids was induced with 100 mM isopropyl-b-D-thiogalactoside (IPTG) separately for each pool (Haas *et al.*, Gene 130:23-21, 1993). Plasmids were transferred into E145 by triparental mating, in which 25 μl of the donor strain (E181), 25 μl of the mobilisator (HB101(pRK2013)), and 50 μl of the recipient strain (E145) were mixed from corresponding bacterial suspensions (O.D.₅₅₀ = 10). The matings were performed for 2-3 hours at 37°C on

-46-

nitrocellulose filters, which were placed on LB plates. Bacteria were suspended in 1 ml LB and aliquots were spread on LB plates containing chloramphenicol, tetracycline, and rifampicin. Each pool gave rise to chloramphenicol-resistant transconjugates in E145, demonstrating that both
5 transposition and conjugation were successful. Generally, several thousand chloramphenicol-resistant transconjugates were obtained, but the number of amp^R colonies varied in different pools, ranging from one to several hundred colonies. Two amp^R colonies from each positive pool were isolated, plasmid DNA was extracted, and the DNA was characterized by further restriction
10 analysis. Only those Tn*Max9* insertions of a single pool that mapped in obviously different plasmid clones, or in markedly different regions of the same clone, were used further.

From 158 of the 198 pools, ampicillin-resistant E145 transconjugates were obtained (80%), showing that in several pools, Tn*Max9* inserted into
15 expressed genes, resulting in production of extracytoplasmic BlaM fusion proteins. Thus, a total of 192 amp^R E145 clones could be isolated by conjugal transfer of plasmids from 198 pools.

To analyze the mutant library, it was determined whether defined gene sequences inactivated by Tn*Max9* were represented once or several times
20 in the whole library. Five transposon-containing plasmids conferring an amp^R phenotype to E145 (pMu7, pMu13, pMu75, pMu94, and pMu110) were randomly selected and DNA fragments flanking the Tn*Max9* insert were isolated and used as probes in Southern hybridization of 120 amp^R clones. The hybridization probes isolated from clones pMu7, pMu75, and pMu94 were
25 between 0.9 and 1.1 kilobases in size, and hybridized exclusively with the inserts of the homologous plasmids. In contrast, the Tn*Max9* flanking regions of clones pMu13 and pMu110 were 4.0 and 5.5 kilobases, respectively. They

-47-

each hybridized with the homologous plasmids, and with one additional clone of the library. Such a result was expected, since the chance of a probe to find a homologous sequence in the library should be higher, the longer the hybridization probes.

5 In order to verify the insertion of the transposon into distinct ORFs encoding putative exported proteins, the *TnMax9*-flanking DNA of five representative *amp^R* mutant clones (pMu7, pMu12, pMu18, pMu20, and pMu26) was sequenced, taking advantage of the M13 forward and reverse primers on *TnMax9* (Fig. 1A). This analysis revealed that the mini-transposon
10 was inserted into different sequences in each plasmid, thereby interrupting ORFs encoding putative proteins. For two clones, the sequences located upstream of the *blaM* gene revealed a putative ribosome-binding site and a potential translational start codon (ATG). Other clones either revealed an ORF spanning the complete sequence (approximately 400 base pairs upstream and
15 downstream of the *TnMax9* insertion) or terminating shortly after the site of *TnMax9* insertion. The partial protein sequences from different ORFs were used for database searches, but no significant homologies with known proteins were found.

 In a further approach, it was determined whether a known gene, like
20 *vacA*, encoding the extracellular vacuolating cytotoxin of *H. pylori*, could be identified using this method and how often such a mutation would be represented in the mutant library. Total cell lysates of the 135 mutants were tested in an immunoblot using the *H. pylori* cytotoxin-specific rabbit antiserum AK197 (Schmitt *et al.*, Mol. Microbiol. 12:307-319, 1994). Two mutants were
25 identified that no longer produced the cytotoxin antigen (mutants P1-26 and P1-47) and partial DNA sequencing of the insertion sites revealed that *TnMax9*

-48-

was inserted at distinct positions in the *vacA* gene, 56 and 53 codons downstream of the ATG start codon.

Thus, the characterization of the mutant collection confirmed that a representative gene library was constructed in *E. coli*, in which target genes encoding exported *H. pylori* proteins were efficiently tagged by TnMax9.

In order to establish a collection of mutants lacking distinct exported proteins, the mutations had to be transferred back into the *H. pylori* chromosome. By means of natural transformation, 86 plasmids could be transformed into the original strain P1. *H. pylori* strains P1 or P12, which were naturally competent for DNA transformation, were transformed with circular plasmid DNA (0.2-0.5 mg/transformation). Transformations to streptomycin resistance were performed with chromosomal DNA (1 mg/transformation), isolated from a streptomycin-resistant NCTC11637 *H. pylori* mutant according to the procedure described in Haas *et al.* (Mol. Microbiol. 8:753-760). Selection was performed on serum plates containing 4 mg/ml chloramphenicol or 500 mg/ml streptomycin. The transformation frequency for a given mutant was calculated as the number of chloramphenicol-, streptomycin-, or erythromycin-resistant colonies per cfu (average of three experiments). The *blaM* gene was deleted by *NotI* digestion, and the plasmid religated, in those plasmids that did not transform strain P1 directly. This procedure, which resulted in a twenty- to thirty-fold higher frequency of transformation, as compared to the same plasmid containing *blaM*, resulted in 36 additional mutant P1 strains. The *blaM*-deletion plasmids that still did not transform strain P1 were used to transform the heterologous *H. pylori* strain P12, possessing an approximately 10-fold higher transformation frequency compared to P1. This resulted in thirteen further mutants.

Thus, from the 192 amp^R plasmids, a total of 135 *H. pylori* mutants (122 mutants in P1 and 13 mutants in P12) were finally obtained by selection for chloramphenicol resistance (70%). The transformation frequency varied between different plasmids in the range of 1×10^{-5} - 1×10^{-7} . The remaining
5 plasmids did not result in any transformants. The collection was frozen as individual mutants in stock cultures at -70°C. To verify the correct insertion of the mini-transposon into the *H. pylori* chromosome, ten representative mutants were tested by Southern hybridization of chromosomal DNA using *cat*_{GC} DNA and the vector pMin2 as probes. Consistent with our previous experience
10 concerning TnMax9-based shuttle mutagenesis of *H. pylori*, the mini-transposon was, in all cases, inserted into the chromosome without integration of the vector DNA, which probably means by a double cross-over, rather than by a single cross-over event. As judged from the hybridization pattern obtained with the *cat* gene as a probe, it appears that TnMax9 is located in different
15 regions of the chromosome, showing that distinct target genes have been interrupted in individual mutants.

The mutants were analyzed for motility, transformation competence, and adherence to KatoIII cells. Screening of the *H. pylori* mutant collection allowed identification of mutants impaired in motility, natural transformation
20 competence, and adherence to gastric epithelial cell lines. Motility mutants could be grouped into distinct classes: (i) mutants lacking the major flagellin subunit FlaA and intact flagella; (ii) mutants with apparently normal flagella, but reduced motility; and (iii) mutants with obviously normal flagella, but completely abolished motility. Two independent mutations, which exhibited
25 defects in natural competence for genetic transformation, mapped to different genetic loci. In addition, two independent mutants were isolated by their failure to bind to the human gastric carcinoma cell line KatoIII. Both mutants

-50-

carried a transposon in the same gene, approximately 0.8 kilobases apart, and showed decrease autoagglutination, when compared to the wild type strain.

Sequences of clones obtained using the above-described transposon shuttle mutagenesis method were used to identify intact genes, lacking inserted
5 transposons, in the *H. pylori* genome, as is described below in Example 5.

The invention is further illustrated by the following examples.
Example 1 describes identification of genes, such as genes that encode the polypeptides of the invention, in the *Helicobacter* genome, as well as identification of signal sequences and primer design for amplification of genes
10 lacking signal sequences. Example 2 describes cloning of DNA encoding GHPO 732, GHPO 419, GHPO 1398, GHPO 706, GHPO 1190, GHPO 986, GHPO 1420, GHPO 1299, and GHPO 13 into a vector that provides a histidine tag, and production and purification of the resulting his-tagged fusion proteins. Example 3 describes methods for cloning DNA encoding the polypeptides of
15 the invention so that they can be produced without his-tags, and Example 4 describes methods for purifying recombinantly produced polypeptides of the invention. Example 5 describes methods for obtaining the nucleic acids of the invention from the deposited clones. Example 6 describes purification of recombinant *H. pylori* antigen GHPO 1190.

20 **EXAMPLE 1: Identification of genes in the *H. pylori* genome, identification of signal sequences, and primer design for amplification of genes lacking signal sequences**

1.A. Creating *H. pylori* genomic databases

The *H. pylori* genome was provided as a text file containing a single
25 contiguous string of nucleotides that had been determined to be 1.76 Megabases in length. The complete genome was split into 17 separate files

-51-

using the program SPLIT (Creativity in Action), giving rise to 16 contigs, each containing 100,000 nucleotides, and a 17th contig containing the remaining 76,000 nucleotides. A header was added to each of the 17 files using the format: >hpg0.txt (representing contig 1), .hpg1.txt (representing contig 2), etc.

5 The resulting 17 files, named hpg0 through hpg16, were then copied together to form one file that represented the plus strand of the complete *H. pylori* genome. The constructed database was given the designation "H." A negative strand database of the *H. pylori* genome was created similarly by first creating a reverse complement of the positive strand using the program SeqPup (D.G.

10 Gilbert, Indiana University Biology Department) and then performing the same procedure as described above for the plus strand. This database was given the designation "N."

The regions predicted to encode open reading frames (ORFs) were defined for the complete *H. pylori* genome using the program GENEMARK™

15 (Borodovsky *et al.*, Comp. Chem. 17:123, 1993). A database was created from a text file containing an annotated version of all ORFs predicted to be encoded by the *H. pylori* genome for both the plus and minus strands, and was given the designation "O." Each ORF was assigned a number indicating its location on the genome and its position relative to other genes. No manipulation of the text

20 file was required.

1.B. Searching the *H. pylori* databases

The databases constructed as is described above were searched using the program FASTA (Pearson *et al.*, Proc. Natl. Acad. Sci. USA 85:2444-2448, 1988). FASTA was used for searching either a DNA sequence against either of

25 the gene databases ("H" and/or "N"), or a peptide sequence against the ORF library ("O"). TFASTX was used to search a peptide sequence against all

-52-

possible reading frames of a DNA database ("H" and/or "N" libraries). Potential frameshifts also being resolved, FASTX was used for searching the translated reading frames of a DNA sequence against either a DNA database, or a peptide sequence against the protein database.

5 1.C. Isolation of DNA sequences from the *H. pylori* genome

The FASTA searches against the constructed DNA databases identified exact nucleotide coordinates on one or more of the isolated contigs, and therefore the location of the target DNA. Once the exact location of the target sequence was known, the contig identified to carry the gene was
10 exported into the software package MapDraw (DNASStar, Inc.) and the gene was isolated. Gene sequences with flanking DNA were then excised and copied into the EditSeq. Software package (DNASStar, Inc.) for further analysis.

1.D. Identification of signal sequences

The deduced protein encoded by a target gene sequence was
15 analyzed using the PROTEAN software package (DNASStar, Inc.). This analysis predicts those areas of the protein that are hydrophobic by using the Kyte-Doolittle algorithm, and identifies any potential polar residues preceding the hydrophobic core region, which is typical for many signal sequences. For confirmation, the target protein was then searched against a PROSITE database
20 (DNASStar, Inc.) consisting of motifs and signatures. Characteristic of many signal sequences and hydrophobic regions in general, is the identification of predicted prokaryotic lipid attachment sites. Where confirmation between the two approaches is apparent at the N-terminus of any protein, putative cleavage sites were sought. Specifically, this includes the presence of either an Alanine
25 (A), Serine (S), or Glycine (G) residue immediately after the core hydrophobic

-53-

region. In the case of lipoproteins, a Cysteine (C) residue would be identified as the +1 residue, post-cleavage.

1.E. Rational design of PCR primers based on the identification of signal sequences

5 To clone gene sequences as N-terminal translational fusions for the generation of recombinant proteins with N-terminal Histidine tags, the gene sequence that specifies the signal sequence is omitted. The 5'-end of the gene-specific portion of the N-terminal primer is designed to start at the first codon beyond the cleavage site. In the case of lipoproteins, the 5'-end of the N-
10 terminal primer begins at the second codon, immediately after the modifiable residue at position +1 post-cleavage. The omission of the signal sequence from the recombinant allows for one-step purification, and potential problems associated with insertion of signal sequences in the membrane of the host strain carrying the hybrid construct are avoided.

15 **EXAMPLE 2: Preparation of isolated DNA encoding GHPO 732, GHPO 419, GHPO 1398, GHPO 706, GHPO 1190, GHPO 986, GHPO 1420, GHPO 1299, and GHPO 13, and production of these polypeptides as histidine-tagged fusion proteins**

2.A. Preparation of genomic DNA from *Helicobacter pylori*

20 *Helicobacter pylori* strain ORV2001, stored in LB medium containing 50% glycerol at -70°C, is grown on Colombia agar containing 7% sheep blood for 48 hours under microaerophilic conditions (8-10% CO₂, 5-7% O₂, 85-87% N₂). Cells are harvested, washed with phosphate buffer saline (PBS) (pH 7.2), and DNA is then extracted from the cells using the Rapid Prep
25 Genomic DNA Isolation kit (Pharmacia Biotech).

2.B. PCR amplification

DNA molecules encoding GHPO 732, GHPO 419, GHPO 1398, GHPO 706, GHPO 1190, GHPO 986, GHPO 1420, GHPO 1299, and GHPO 13 are amplified from genomic DNA, as can be prepared as is described above, by the Polymerase Chain Reaction (PCR) using the following primers:

GHPO 732 (HPO 64):

N-terminal primer:

5'-GCCGGATCCATGACTTATGGGTATGGGGAA-3' (SEQ ID NO:171);

and

10 C-terminal primer:

5'-GCCCTCGAGACTTTTATTGATTCACCATTTTCATT-3' (SEQ ID NO:172).

GHPO 419 (HPO 54):

N-terminal primer:

15 5'-GCCGGATCCATCGCTGAAGAAAATGGGGCG-3' (SEQ ID NO:173);

and

C-terminal primer:

5'-GCCCCGGCCGCCCTAAAACTATAAACATAACTC-3' (SEQ ID NO:174).

20 GHPO 1398 (HPO 15):

N-terminal primer:

5'-GCCGGATCCGGTATTAGGAAGCTTATACCATC-3' (SEQ ID NO:175);

and

C-terminal primer:

25 5'-GCCCTCGAGAAGTTCTATTTTAAATTCCTTGAGAG-3' (SEQ ID NO:176).

GHPO 706 (HPO 50):

-55-

N-terminal primer:

5'-GCCGGATCCTCTGATAGCCATAAAGAAAAAAGGAC-3' (SEQ ID

NO:177); and

C-terminal primer:

5 5'-GCCCTCGAGATCTTTAGAAATCAACCCCCAAAGC-3' (SEQ ID

NO:178).

GHPO 1190 (HPO 76):

N-terminal primer:

5'-GCCGGATCCGACTTAGAACATTTTAACACGCTC-3' (SEQ ID

10 NO:179);

and

C-terminal primer:

5'-GCCCTCGAGTCATTTTAAACGACTCAAAACAAA-3' (SEQ ID

NO:180).

15 GHPO 986:

N-terminal primer:

5'-GCCGGATCCGGCCAAAGCGTGCGCACTTATTGG-3' (SEQ ID

NO:181); and

C-terminal primer:

20 5'-GCCCTCGAGTTATTGTTCCAACCCCCACGCATC-3' (SEQ ID NO:182).

GHPO 1420:

N-terminal primer:

5'-GCCGGATCCAAGAGCAATGCTGATGACAAACC-3' (SEQ ID

NO:183); and

25 C-terminal primer:

5'-GCCCTCGAGTTATGAGTTAAAGCCCCTTGTC-3' (SEQ ID NO:184).

GHPO 1299:

-56-

N-terminal primer:

5'-GCCGGATCCGAATCAGTAAAAACAGGAAAAAC-3' (SEQ ID NO:185); and

C-terminal primer:

5'-GCCCTCGAGCGGCTCTTTGGAGTTTATTG-3' (SEQ ID NO:186).

GHPO 13:

N-terminal primer:

5'-GCCGGATCCATCATTCCCTCTCGCTCTATGG-3' (SEQ ID NO:187);
and

C-terminal primer:

5'-GCCCTCGAGACCTTAATGCGTTGCGTTTTCTTT-3' (SEQ ID NO:188).

The N-terminal and C-terminal primers for each clone both include a 5' clamp and a restriction enzyme recognition sequence for cloning purposes (*Bam*HI (GGATCC) and *Xho*I (CTCGAG) or *Not*I (CGGCCG) recognition sequences). The N-terminal primer is designed so that the amplified product does not encode the signal sequence and the potential cleavage site.

Amplification of gene-specific DNA is carried out using Pwo DNA Polymerase (Boehringer Mannheim), which is a proof-reading polymerase, according to general guidance provided by the manufacturer. Because of the exonuclease activity of the polymerase, two reaction mixtures (mixtures 1 and 2) are first prepared separately and combined just prior to amplification. These mixtures are as follows:

-57-

<u>Ingredient (final conc.)</u>		<u>Mixture 1 (μl)</u>	<u>Mixture 2 (μl)</u>
5	distilled H ₂ O	160	79
	dNTPs (200 μ M each)	40	---
	10x PCR buffer	---	20
	primers (100 nM each)	1	---
	DNA template (200 ng) as obtained in 5.A.	2	---
(10x PCR buffer contains 100 mM Tris-HCl (pH 8.85), 250 mM KCl, 50 mM (NH ₄) ₂ SO ₄ , 20 mM MgSO ₄)			

10 Amplification is carried out as follows:

<u>Cycling conditions</u>	<u>Temp ($^{\circ}$C)</u>	<u>Time (min.)</u>	<u>Number of cycles</u>
Initial denaturing step	96	4	1
Denaturing step	94	0.5	20
Annealing step	50	1	20
15 Extension step	72	1	20
Final extension step	72	5	1

2.C. Transformation and selection of transformants

A single PCR product is thus amplified and is then digested at 37 $^{\circ}$ C for 2 hours with *Bam*HI and *Xho*I or *Not*I concurrently in a 20 μ l reaction volume. The digested product is ligated to similarly cleaved pET28a (Novagen) that is dephosphorylated prior to the ligation by treatment with Calf Intestinal Alkaline Phosphatase (CIP). The gene fusion constructed in this manner allows one-step affinity purification of the resulting fusion protein because of the presence of histidine residues at the N-terminus of the fusion protein, which are encoded by the vector.

The ligation reaction (20 μ l) is carried out at 14 $^{\circ}$ C overnight and then is used to transform 100 μ l fresh *E. coli* XL1-blue competent cells (Novagen). The cells are incubated on ice for 2 hours, heat-shocked at 42 $^{\circ}$ C for 30 seconds, and returned to ice for 90 seconds. The samples are then added to 1 ml LB broth in the absence of selection and grown at 37 $^{\circ}$ C for 2 hours.

-58-

The cells are plated out on LB agar containing kanamycin (50 $\mu\text{g/ml}$) at a 10x and neat dilution and incubated overnight at 37°C. The following day, 50 colonies are picked onto secondary plates and incubated at 37°C overnight.

Five colonies are picked into 3 ml LB broth supplemented with
5 kanamycin (100 $\mu\text{g/ml}$) and are grown overnight at 37°C. Plasmid DNA is extracted using the Quiagen mini-prep. method and is quantitated by agarose gel electrophoresis.

PCR is performed with the gene-specific primers under the conditions set forth above and transformant DNA is confirmed to contain the
10 desired insert. If PCR-positive, one of the five plasmid DNA samples (500 ng) extracted from the *E. coli* XL1-blue cells is used to transform competent BL21 (λDE3) *E. coli* competent cells (Novagen; as described previously). Transformants (10) are picked onto selective kanamycin (50 $\mu\text{g/mL}$) containing LB agar plates and stored as a research stock in LB containing 50% glycerol.

15 2.D. Purification of recombinant proteins

One ml of frozen glycerol stock prepared as described in 2.C. is used to inoculate 50 ml of LB medium containing 25 $\mu\text{g/ml}$ of kanamycin in a 250 ml Erlenmeyer flask. The flask is incubated at 37°C for 2 hours or until the absorbance at 600 nm (OD_{600}) reaches 0.4-1.0. The culture is stopped from
20 growing by placing the flask at 4°C overnight. The following day, 10 ml of the overnight culture are used to inoculate 240 ml LB medium containing kanamycin (25 $\mu\text{g/ml}$), with the initial OD_{600} about 0.02-0.04. Four flasks are inoculated for each ORF. The cells are grown to an OD_{600} of 1.0 (about 2 hours at 37°C), a 1 ml sample is harvested by centrifugation, and the sample is
25 analyzed by SDS-PAGE to detect any leaky expression. The remaining culture

-59-

is induced with 1 mM IPTG and the induced cultures are grown for an additional 2 hours at 37°C.

The final OD₆₀₀ is taken and the cells are harvested by centrifugation at 5,000 x g for 15 minutes at 4°C. The supernatant is discarded and the pellets
5 are resuspended in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA. Two hundred and fifty ml of buffer are used for 1 liter of culture and the cells are recovered by centrifugation at 12,000 x g for 20 minutes. The supernatant is discarded and the pellets are stored at -45°C.

2. E. Protein purification

10 Pellets obtained from 2.D. are thawed and resuspended in 95 ml of 50 mM Tris-HCl (pH 8.0). Pefabloc and lysozyme are added to final concentrations of 100 µM and 100 µg/ml, respectively. The mixture is homogenized with magnetic stirring at 5°C for 30 minutes. Benzonase (Merck)
15 is added at a 1 U/ml final concentration, in the presence of 10 mM MgCl₂, to ensure total digestion of the DNA. The suspension is sonicated (Branson Sonifier 450) for 3 cycles of 2 minutes each at maximum output. The homogenate is centrifuged at 19,000 x g for 15 minutes and both the supernatant and the pellet are analyzed by SDS-PAGE to detect the cellular
20 location of the target protein in the soluble or insoluble fractions, as is described further below.

2.E.1. Soluble fraction

If the target protein is produced in a soluble form (*i.e.*, in the supernatant obtained in 2.E.) NaCl and imidazole are added to the supernatant to final concentrations of 50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, and 10 mM
25 imidazole (buffer A). The mixture is filtered through a 0.45 µm membrane and

-60-

loaded onto an IMAC column (Pharmacia HiTrap chelating Sepharose; 1 ml), which has been charged with nickel ions according to the manufacturer's recommendations. After loading, the column is washed with 50 column volumes of buffer A and the recombinant target protein is eluted with 5 ml of
5 buffer B (50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 500 mM imidazole).

The elution profile is monitored by measuring the absorbance of the fractions at 280 nm. Fractions corresponding to the protein peak are pooled, dialyzed against PBS containing 0.5 M arginine, filtered through a 0.22 µm membrane, and stored at -45°C.

10 2.E.2. Insoluble fraction

If the target protein is expressed in the insoluble fraction (pellets obtained from 2.E.), purification is conducted under denaturing conditions. NaCl, imidazole, and urea are added to the resuspended pellet to final concentrations of 50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 10 mM imidazole,
15 and 6 M urea (buffer C). After complete solubilization, the mixture is filtered through a 0.45 µm membrane and loaded onto an IMAC column.

The purification procedures on the IMAC column are the same as described in 2.E.1., except that 6 M urea is included in all buffers used and 10 column volumes of buffer C are used to wash the column after protein loading,
20 instead of 50 column volumes.

The protein fractions eluted from the IMAC column with buffer D (buffer C containing 500 mM imidazole) are pooled. Arginine is added to the solution to final concentration of 0.5 M and the mixture is dialyzed against PBS containing 0.5 M arginine and various concentrations of urea (4 M, 3 M, 2 M, 1
25 M, and 0.5 M) to progressively decrease the concentration of urea. The final dialysate is filtered through a 0.22 µm membrane and stored at -45°C.

-61-

Alternatively, when the above purification process is not as efficient as it should be, two other processes may be used as follows. A first alternative involves the use of a mild denaturant, N-octyl glucoside (NOG). Briefly, a pellet obtained in 2.E. is homogenized in 5 mM imidazole, 500 mM sodium chloride, 20 mM Tris-HCl (pH 7.9) by microfluidization at a pressure of 15,000 psi and is clarified by centrifugation at 4,000-5,000 x g. The pellet is recovered, resuspended in 50 mM NaPO₄ (pH 7.5) containing 1-2 % weight /volume NOG, and homogenized. The NOG-soluble impurities are removed by centrifugation. The pellet is extracted once more by repeating the preceding extraction step. The pellet is dissolved in 8 M urea, 50 mM Tris (pH 8.0). The urea-solubilized protein is diluted with an equal volume of 2 M arginine, 50 mM Tris (pH 8.0), and is dialyzed against 1 M arginine for 24-48 hours to remove the urea. The final dialysate is filtered through a 0.22 µm membrane and stored at -45°C.

A second alternative involves the use of a strong denaturant, such as guanidine hydrochloride. Briefly, a pellet obtained in 2.E. is homogenized in 5 mM imidazole, 500 mM sodium chloride, 20 mM Tris-HCl (pH 7.9) by microfluidization at a pressure of 15,000 psi and clarified by centrifugation at 4,000-5,000 x g. The pellet is recovered, resuspended in 6 M guanidine hydrochloride, and passed through an IMAC column charged with Ni⁺⁺. The bound antigen is eluted with 8 M urea (pH 8.5). Beta-mercaptoethanol is added to the eluted protein to a final concentration of 1 mM, then the eluted protein is passed through a Sephadex G-25 column equilibrated in 0.1 M acetic acid. Protein eluted from the column is slowly added to 4 volumes of 50 mM phosphate buffer (pH 7.0). The protein remains in solution.

-62-

2.F. Evaluation of the protective activity of the purified protein

Groups of 10 Swiss Webster mice (Taconic Labs) are immunized rectally with 25 μ g of the purified recombinant protein, admixed with 1 μ g of cholera toxin (Berna) in physiological buffer. Mice are immunized on days 0, 5 7, 14, and 21. Fourteen days after the last immunization, the mice are challenged with *H. pylori* strain ORV2001 grown in liquid media (the cells are grown on agar plates, as described in 2.A., and, after harvest, the cells are resuspended in Brucella broth; the flasks are then incubated overnight at 37°C). Fourteen days after challenge, the mice are sacrificed and their stomachs are 10 removed. The amount of *H. pylori* is determined by measuring the urease activity in the stomach and by culture.

2.G. Production of monospecific polyclonal antibodies**2.G.1. Hyperimmune rabbit antiserum**

New Zealand rabbits are injected both subcutaneously and 15 intramuscularly with 100 μ g of a purified fusion polypeptide, as obtained in 2.E.1. or 2.E.2., in the presence of Freund's complete adjuvant and in a total volume of approximately 2 ml. Twenty one and 42 days after the initial injection, booster doses, which are identical to priming doses, except that Freund's incomplete adjuvant is used, are administered in the same way. 20 Fifteen days after the last injection, animal serum is recovered, decomplemented, and filtered through a 0.45 μ m membrane.

2.G.2. Mouse hyperimmune ascites fluid

Ten mice are injected subcutaneously with 10-50 μ g of a purified fusion polypeptide as obtained in 2.E.1. or 2.E.2., in the presence of Freund's complete 25 adjuvant and in a volume of approximately 200 μ l. Seven and 14 days after the

-63-

initial injection, booster doses, which are identical to the priming doses, except that Freund's incomplete adjuvant is used, are administered in the same way. Twenty one and 28 days after the initial infection, mice receive 50 μ g of the antigen alone intraperitoneally. On day 21, mice are also injected

5 intraperitoneally with sarcoma 180/TG cells CM26684 (Lennette *et al.*, *Diagnostic Procedures for Viral, Rickettsial, and Chlamydial Infections*, 5th Ed. Washington DC, American Public Health Association, 1979). Ascites fluid is collected 10-13 days after the last injection.

10 **EXAMPLE 3: Methods for producing transcriptional fusions lacking His-tags**

Methods for amplification and cloning of DNA encoding the polypeptides of the invention as transcriptional fusions lacking His-tags are described as follows. Two PCR primers for each clone are designed based upon the sequences of the polynucleotides that encode them (SEQ ID NOs: 1-15 169 (odd numbers)). These primers can be used to amplify DNA encoding the polypeptides of the invention from any *Helicobacter pylori* strain, including, for example, ORV2001 and the strain deposited as ATCC deposit number 43579, as well as from other *Helicobacter* species.

The N-terminal primers are designed to include the ribosome binding site of the target gene, the ATG start site, and any signal sequence and cleavage site. The N-terminal primers can include a 5' clamp and a restriction endonuclease recognition site, such as that for *Bam*HI (GGATCC), which facilitates subsequent cloning. Similarly, the C-terminal primers can include a restriction endonuclease recognition site, such as that for *Xho*I (CTCGAG),

25 which can be used in subsequent cloning, and a TAA stop codon.

-64-

Amplification of genes encoding the polypeptides of the invention is carried out using Thermalase DNA Polymerase under the conditions described above in Example 2. Alternatively, Vent DNA polymerase (New England Biolabs), Pwo DNA polymerase (Boehringer Mannheim), or Taq DNA
5 polymerase (Appligene) can be used, according to instructions provided by the manufacturers.

A single PCR product for each clone is amplified and cloned into appropriately cleaved pET 24 (*e.g.*, *Bam*HI-*Xho*I cleaved pET 24), resulting in construction of a transcriptional fusion that permits expression of the proteins
10 without His-tags. The expressed products can be purified as denatured proteins that are refolded by dialysis into 1 M arginine.

Cloning into pET 24 allows transcription of the genes from the T7 promoter, which is supplied by the vector, but relies upon binding of the RNA-specific DNA polymerase to the intrinsic ribosome binding sites of the genes,
15 and thereby expression of the complete ORF. The amplification, digestion, and cloning protocols are as described above for constructing translational fusions.

Amplification of clone GHPO 1190 DNA

Design of PCR primers for cloning

Two PCR primers are designed based on the complete gene sequence
20 (see table 1). The N-terminal primer (FC1) is designed to include the ribosome binding site of the target gene, the ATG start site, and the signal sequence (with cleavage site). It includes a clamp (GCC) at the 5' most end, and a *Sac*I recognition sequence (GAGCTC) for cloning purposes.

The C-terminal primer (RN2) includes an *Xho*I recognition sequence for
25 cloning purposes, and the natural TAA stop codon.

N-terminal primer (FC1):

-65-

5'-GCCGAGCTCCAAGCAAAAAAATGTCAATTAAAAGGG-3' (SEQ ID NO:189)

C-terminal primer (RN2):

5'-GCCCTCGAGGTCTAAATTAGAATAAGTGTTGTT-3' (SEQ ID NO:190)

5 Amplification of each specified gene can be achieved by employing FC1/RN2 primers for any of the genes described (see Table 1).

PCR conditions

Amplification of gene-specific DNA is carried out using Pwo DNA Polymerase (Boehringer Mannheim) under the following conditions. Due to the exonuclease activity of the polymerase, two reaction mixtures are prepared separately and combined just prior to amplification.

Reaction ingredients:	<u>Ingredient (final conc.)</u>	<u>Mixture 1 (μl)</u>	<u>Mixture 2 (μl)</u>
15	distilled H ₂ O	160	79
	dNTPs (200 μM each)	40	-
	10X buffer	-	20
	primer 1 (100 nM)	1	-
	primer 2 (100 nM)	1	-
	Template (200 ng)	2	0
20	<u>Cycling condition</u>	<u>Temp (°C)</u>	<u>Time(min)</u>
	Initial denaturing step	96	4
	Denaturing step	94	0.5
	Annealing step	50	1
	Extension step	72	1
	Final extension step	72	1

25 A single PCR product of 624 basepairs is amplified and cloned into *SacI-XhoI* cleaved pET 24, allowing construction of a transcriptional fusion and expression of GHPO 1190 antigen in the absence of a His-tag. In this instance, expressed product can be purified as a denatured protein that is re-folded by dialysis into 1 M arginine.

-66-

Cloning into pET 24 allows transcription from the T7 promoter, supplied by the vector, but relies upon binding of the RNA-specific DNA polymerase to the intrinsic ribosome binding site for GHPO 1190, and thereby expression of the complete ORF. The amplification, restriction, and cloning
5 protocols are as previously described for constructing translational fusions.

EXAMPLE 4: Purification of the polypeptides of the invention by immunoaffinity

4.A. Purification of specific IgGs

An immune serum, as prepared in section 2.G., is applied to a protein A
10 Sepharose Fast Flow column (Pharmacia) equilibrated in 100 mM Tris-HCl (pH 8.0). The resin is washed by applying 10 column volumes of 100 mM Tris-HCl and 10 volumes of 10 mM Tris-HCl (pH 8.0) to the column. IgG antibodies are eluted with 0.1 M glycine buffer (pH 3.0) and are collected as 5 ml fractions to which is added 0.25 ml 1 M Tris-HCl (pH 8.0). The optical
15 density of the eluate is measured at 280 nm and the fractions containing the IgG antibodies are pooled, dialyzed against 50 mM Tris-HCl (pH 8.0), and, if necessary, stored frozen at
-70°C.

4.B. Preparation of the column

20 An appropriate amount of CNBr-activated Sepharose 4B gel (1 g of dried gel provides for approximately 3.5 ml of hydrated gel; gel capacity is from 5 to 10 mg coupled IgG/ml of gel) manufactured by Pharmacia (17-0430-01) is suspended in 1 mM HCl buffer and washed with a buchner by adding small quantities of 1 mM HCl buffer. The total volume of buffer is 200 ml per
25 gram of gel.

-67-

Purified IgG antibodies are dialyzed for 4 hours at $20 \pm 5^\circ\text{C}$ against 50 volumes of 500 mM sodium phosphate buffer (pH 7.5). The antibodies are then diluted in 500 mM phosphate buffer (pH 7.5) to a final concentration of 3 mg/ml.

5 IgG antibodies are mixed with the gel overnight at $5 \pm 3^\circ\text{C}$. The gel is packed into a chromatography column and is washed with 2 column volumes of 500 mM phosphate buffer (pH 7.5), and 1 column volume of 50 mM sodium phosphate buffer, containing 500 mM NaCl (pH 7.5). The gel is then transferred to a tube, mixed with 100 mM ethanolamine (pH 7.5) for 4 hours at
10 room temperature, and washed twice with 2 column volumes of PBS. The gel is then stored in 1/10,000 PBS/merthiolate. The amount of IgG antibodies coupled to the gel is determined by measuring the optical density (OD) at 280 nm of the IgG solution and the direct eluate, plus washings.

4.C. Adsorption and elution of the antigen

15 An antigen solution in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, for example, the supernatant obtained in 3.E. or the solubilized pellet obtained in 3.E., after centrifugation and filtration through a $0.45 \mu\text{m}$ membrane, is applied to a column equilibrated with 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, at a flow rate of about 10 ml/hour. The column is then washed with 20 volumes of
20 50 mM Tris-HCl (pH 8.0), 2 mM EDTA. Alternatively, adsorption can be achieved by mixing overnight at $5 \pm 3^\circ\text{C}$.

The adsorbed gel is washed with 2 to 6 volumes of 10 mM sodium phosphate buffer (pH 6.8) and the antigen is eluted with 100 mM glycine buffer (pH 2.5). The eluate is recovered in 3 mL fractions, to each of which is added
25 $150 \mu\text{l}$ of 1 M sodium phosphate buffer (pH 8.0). Absorption is measured at

-68-

280 nm for each fraction; those fractions containing the antigen are pooled and stored at -20°C.

EXAMPLE 5: Preparation of isolated DNA encoding the polypeptides of the invention from the deposited clones.

As mentioned above, *E. coli* strains including plasmids containing nucleic acids encoding GHPO 1190 (formerly HPO76, ATCC# 98197), GHPO 1212 (formerly HPO18, ATCC# 98210), GHPO 1012 (formerly HPO121, ATCC# 98201), GHPO 1501 (formerly HPO45, ATCC# 98208), GHPO 1688 (formerly HPO101, ATCC# 98198), GHPO 346 (formerly HPO116, ATCC# 98200), GHPO 1200 (formerly HPO7, ATCC# 98211), GHPO 1538 (formerly HPO104, ATCC# 98199), GHPO 1398 (formerly HPO15, ATCC# 98214), GHPO 1001 (formerly HPO58, ATCC# 98206), GHPO 470 (formerly HPO132, ATCC# 98202), GHPO 689 (formerly HPO9, ATCC# 98203), GHPO 1550 (formerly HPO38, ATCC# 98204), GHPO 1620 (formerly HPO87, ATCC# 98205), GHPO 574 (formerly HPO71, ATCC# 98217), GHPO 329 (formerly HPO70, ATCC# 98219), GHPO 1374 (formerly HPO80, ATCC# 98215), GHPO 956 (formerly HPO95 ATCC# 98216), HPO 98 (ATCC# 98218), GHPO 1346 (formerly HPO57, ATCC# 98220), GHPO 706 (formerly HPO50, ATCC# 98207), GHPO 732 (formerly HPO64, ATCC# 98213), GHPO 419 (formerly HPO54, ATCC# 98212), and GHPO 276 (formerly HPO42, ATCC# 98209) were deposited in *E. coli* strain DH5α under the Budapest Treaty with the American Type Culture Collection (ATCC; Rockville, Maryland) on October 9, 1996 and were designated with accession numbers indicated in parentheses above. These plasmids each contain a genomic DNA *Bgl*II-*Cla*I insert from *H. pylori* strain P1 or P12 (referred to as

-69-

69-A and 888-0 in Haas *et al.*, Mol. Microbiol. (1993) 8:753). Each of the inserts are disrupted by the presence of transposon TnMax9 (Kahrs *et al.*, Gene (1995) 167:53). DNA molecules lacking the transposon can be amplified from the plasmids using standard PCR techniques, such as inverse and recombinant
5 PCR (see, *e.g.*, Innis *et al.*, *supra*), so that a full-length *H. pylori* insert is reconstituted. For example, the *H. pylori* sequences flanking the transposon can each be amplified by PCR, and then ligated together to form the full-length *H. pylori* gene lacking the transposon. Primers that can be used in these methods for each of the twenty-four deposited clones of the invention are
10 shown in Table 1. The locations of insertion of the transposon in each of the deposited clones are between the nucleotides indicated in parentheses after the name of each clone, as follows: HPO101 (497-498), GHPO 1538 (428-429), GHPO 346 (433-444), GHPO 1012 (463-464), GHPO 132 (408-409), GHPO 1212 (226-227), GHPO 1550 (347-348), GHPO 276 (372-373), GHPO 1501
15 (299-300), GHPO 706 (29-293), GHPO 419 (351-352), GHPO 1346 (266-267), GHPO 1001 (434-435), GHPO 732 (224-225), GHPO 329 (114-115), GHPO 574 (274-275), GHPO 1190 (412-413), GHPO 1200 (349-350), GHPO 1374 (105-106), GHPO 1620 (26-27), GHPO 956 (64-65), HPO 98 (43-44), and GHPO 689 (346-347).

20 **EXAMPLE 6: Purification of recombinant *H. pylori* antigen from GHPO 1190.**

A pellet of *E. coli* expressing GHPO 1190 is homogenized in 5 mM imidazole, 500 mM sodium chloride, 20 mM Tris-HCl (pH 7.9) by microfluidization at a pressure of 15,000 psi, and clarified by centrifugation at
25 4000-5000g.

Method 1

-70-

The pellet containing cloned protein is suspended in buffer containing 2% N-octyl glucoside (NOG) and is homogenized. The NOG soluble protein is removed by centrifugation. The pellet is extracted one more time with 2% NOG. After centrifugation, the pellet is dissolved in 8 M urea. The urea solubilized protein is diluted with an equal volume of 2 M arginine and dialyzed against 1 M arginine for 24-48 hours to remove urea. The cloned protein remains in solution. SDS-PAGE and Coomassie staining, followed by densitometric scanning, shows that the protein is 80-85% pure cloned antigen.

Method 2

The pellet containing cloned protein is solubilized in 6 M guanidine hydrochloride and is passed through an IMAC column charged with Ni^{++} . The bound antigen is eluted with 8 M urea (pH 8.5). β -Mercaptoethanol is added to eluted protein to a final concentration of 1 mM, then passed through a Sephadex G-25 column equilibrated in 0.1 M acetic acid. Protein eluted from Sephadex G-25 column is slowly added to 4 volumes of 50 mM phosphate (pH 7.0). The protein remains in solution.

Purification of recombinant proteins

Recombinant proteins expressed as Histidine-tagged fusion proteins can be solubilized and purified by using a metal affinity column (nickel column). The bound protein can be eluted with imidazole buffer, with or without urea, or by using low pH buffers, with or without urea. Urea or guanidine hydrochloride-denatured proteins can then be renatured using appropriate renaturing buffers. With a number of recombinant *H. pylori* antigens (HpaA and clone GHPO 1190), renaturation conditions using arginine hydrochloride (0.25-1 M) have been determined.

-71-

Recombinant proteins without a His-tag can be solubilized and purified using immunoaffinity, ion-exchange, sizing, and/or hydrophobic chromatography. Proteins expressed as insoluble aggregates in inclusion bodies can be solubilized in denaturing agents, such as 8 M urea or 6 M
5 guanidine hydrochloride. Appropriate folding and renaturation can readily be determined by one skilled in the art.

The above pellet containing cloned protein is suspended in 50 mM NaPO₄ (pH 7.5) containing 1% weight/volume N-octyl glucoside (NOG) and mixed vigorously. The NOG soluble impurities are removed by centrifugation.
10 The remaining pellet is extracted one more time with the 1% NOG solution to further remove impurities. After centrifugation, the pellet is solubilized in 8 M urea, 50 mM Tris (pH 8.0). The Urea solubilized protein is diluted with an equal volume of 2 M Arginine, 50 mM Tris (pH 8.0), and is dialyzed against 1
15 M Arginine, 50 mM Tris, 50 mM NaCl (pH 8.0) for 24-48 hours to remove urea. The cloned protein remains in solution following dialysis. SDS-PAGE and Coomassie staining followed by densitometric scanning shows that the protein is 80-85% pure cloned antigen.

Other embodiments are within the following claims.

RE-CONSTRUCTION OF A COMPLETE ORF BY RECOMBINANT PCR

F' denotes forward primer
 R' denotes reverse primer
 C' denotes coding strand
 N' denotes non-coding strand

All FC1 and RN2 primers have incorporated at their 5' end a clamp and a recognition sequence for cloning purposes
 GGC clamp present for amplification and cloning of entire gene sequence from chromosomal DNA

[X] denotes any nucleotide sequence not present in the completed gene sequence

() Identifies region of overlap between the two original PCR products, and is consistently 10 nucleotides long for each clone

CLONE No.	Primer type	nt positions	Primer sequence (5' - 3')	Length of gene seq.	T _m (oC)
76	FC1	304 - 330	GCC [X] CAAGCAAAAAATGCAATTTAAAGGG	27	70
	RN1	413 - 391	TAGTCCATACGATAGCCTATG	22	62
	FC2	404 - 436	(TATGGAACCTA) GAACATTTTAAACACGCTCTATTA	33	60
	RN2	927 - 904	GCC [X] GTCTAAATTAGAATAAGTGTGTT	24	60
18	FC1	101 - 124	GCC [X] AATATATGGGAACCTTAATGAGAAT	24	60
	RN1	227 - 206	TGCGAGATTAACTGTTTCA	22	60
	FC2	218 - 249	(AAATCTCGCA) GAAATCTTTCACACGCGAGCAA	32	60
	RN2	922 - 901	GCC [X] ATGTCATGTCAAACTATGAAGC	22	60
121	FC1	141 - 164	GCC [X] TCACAATGGATAAAACAACAACA	24	62
	RN1	451 - 473	GCCCTTTTGTAGGGGTAG	21	62
	FC2	455 - 485	(ACAAAAGGGC) TTTTAGAGCATGTGAGCCATC	32	62
	RN2	814 - 796	GCC [X] CTGTCCAAATCAGCCACCC	19	60
45	FC1	1 - 26	GCC [X] ATGAAAAGATTGATTTGTTTTATC	26	62
	RN1	299 - 278	AGCCGTATTGTTGTTTGCG	22	62
	FC2	290 - 323	(AATACGGCTTTAAAGCTATAGAAAATTTAAACGC)	34	60
	RN2	603 - 582	GCC [X] TTAATATCCCAATCCTGCCAC	22	62

101	FC1	308 - 332	GCC [X] GAAGGATTTATTATGATTAAAGAA	25	60
	RN1	497 - 474	AACCTAATTTGAAATTCAAACCAT	24	60
	FC2	488 - 519	(AAATTAGTT) TTGAGGCTTTGCCAATAATG	32	60
	RN2	893 - 869	GCC [X] AAGGAATAAATTAGAAAGTGAAGAA	25	62
116	FC1	236 - 259	GCC [X] CGCATTTGATTGATGAATAAMCC	23	62
	RN1	434 - 416	CGCCTATAACCGCTCCATT	19	60
	FC2	425 - 456	(GTTATAGCG) ATAAAGGTTTAAACGCAGCTAAG	32	60
	RN2	812 - 790	GCC [X] CTCACTAAMAGCAATTTTGTGAG	23	60
7	FC1	195 - 220	GCC [X] TAAGGAATGAAGTTGATAAAATTTGT	26	64
	RN1	349 - 327	GCAATTTTCATTCCTTTGGAC	23	60
	FC2	339 - 371	(ATGAAATGC) ACGCCCAATAATAAGGAAGTA	32	60
	RN2	738 - 717	GCC [X] GGATTTATTGAGCTTCCCCCTT	22	62
104	FC1	251 - 271	GCC [X] AAGGGCGAATGAGCAAGA	21	60
	RN1	429 - 407	TAAATAACCAACAGAGTGATCA	23	60
	FC2	420 - 452	(GGTTATTTT) GTGGATATTTGGGTTTATAGCGA	33	62
	RN2	784 - 761	GCC [X] TTTTAAAGAAATCACTTCTCTCGG	24	62
58	FC1	118 - 143	GCC [X] ATAGGAACAAGCATGTTTTTAAAC	26	66
	RN1	434 - 413	TGAAGTCTTGGGATTTTGTCTT	22	60
	FC2	425 - 454	(CAAGACTTCA) AAAAAGAGGAGCGGTTGCC	30	60
	RN2	650 - 630	GCC [X] CTGGCTTATTGCGTATCATC	20	60
132	FC1	294 - 314	GGC [X] GGAAGAATAATGCTCGCTTCC	21	62
	RN1	409 - 378	ACTGGAGTGTGGATAAAACTAT	22	60
	FC2	400 - 430	(ACACTCCAGT) AGATGCTTCCCGGATATTTT	31	60
	RN2	761 - 741	GCC [X] CTATCTCCAGGGATATGGCC	21	64
9	FC1	211 - 233	GCC [X] GATGGATTTTTATGGGGGTGAG	23	64
	RN1	347 - 328	GGCACTGCCGCAGATTCTA	19	60
	FC2	338 - 370	(CGGCAGTGC) TTAGCCTATTATTAGAAGCGA	33	60
	RN2	686 - 665	GCC [X] ATGGTATTTGCTAAGACCCCTC	22	62

38	FC1	220 - 242	GCC [X] AAAAGGGTTTTAAATATGGCTG	23	60
	RN1	348 - 327	ACAAGGATAAAAACGCGCTAA	22	60
	FC2	239 - 371	(TTATCCTTG) TGC TGGCTGGTTTTTTTAAATT	33	60
	RN2	597 - 575	GCC [X] AAGATTCTAAAGGGCTTCAAAAT	23	60
71	FC1	1 - 25	GCC [X] ATGTTGAAATTTAAATATGGTTTGA	25	60
	RN1	274 - 254	AAACCCCACTCTTATCATCGG	21	62
	FC2	265 - 294	(AGTGGGGTT) TTTTAGGGGGTGGGTATGCT	30	60
	RN2	524 - 505	GCC [X] GAGCCTACAGGTTGCTTGC	20	60
70	FC1	1 - 23	GCC [X] ATGGTATTTGACAGAACAAATCAG	23	62
	RN1	115 - 96	GAAAGCCACCCCGCTTATT	20	60
	FC2	106 - 137	(GTGGCTTTT) AAAAGAGTGGTGGCAACAATT	32	60
	RN2	495 - 471	GCC [X] TTAGGAATAGCATAACAACAACACG	25	66
80	FC1	1 - 25	GCC [X] ATGTTAGAAAAATTGAATTGAAAGAG	25	62
	RN1	106 - 95	TGAACACATAGCCTAAACCCAC	21	62
	FC2	97 - 127	(TATGTGTTCA) TGAAGAGTTGTGGCACATGC	31	62
	RN2	435 - 415	GCC [X] TTATGCGATAGGGGCGGTATC	21	66
95	FC1	1 - 27	GCC [X] ATGAAAAATTTTTTCTCAATCTTT	27	60
	RN1	64 - 46	TGGCCAGTAGCGGTTTCAT	19	60
	FC2	55 - 98	(CTACTGGCCA) TGGATGGCAATGGCGTTTTTTTAG	34	68
	RN2	432 - 408	GCC [X] TTATTGATGAACATTAACCAATTA	25	60
98	FC1	1 - 22	GCC [X] ATGAAAAACCTTTAAAAACCTGC	22	58
	RN1	43 - 23	TAGCGATCAGGCTAAACAGA	21	60
	FC2	34 - 62	(CTGATCGTA) TGAGTTGGCTCCAAGCGGA	29	60
	RN2	336 - 313	GCC [X] TTAAAACTCATAGCGTTTTTTCAAT	24	60
42	FC1	18-51	GCC [X] GAGAGTAGTGGCAGAGTTTATGCTGATCC	34	98
	RN1	380-351	(AACTTTTC) TCTATCCCAATTCGTTACGCTC	30	64
	FC2	366-396	(GGATAGA) GAAAAGTTTGGCGTCAAAAGTTGG	31	68
	RN2	822-801	GCC [X] GGCTTAAACTGGAACGGATTTC	22	64

SUBSTITUTE SHEET (RULE 26)

50	FC1 RN1 FC2 RN2	140-170 297-270 287-314 607-584	GCC [X] TAAAGTTTGTCTAAAGATGGTTTAATT (GACTTCTAAAG)CGTCCTTTTTCCTTA (CTTTA)GAAGTCATTAAACAAGAGGGGT GCC [X] CCCATCTTTAGAAATCAACCCCA	31 28 29 24	76 56 64 70
64	FC1 RN1 FC2 RN2	23-50 225-149 216-244 1039-1012	GCC [X] GAAATCAAGGAGTTTGTATGCAACAGCG (A)AGCTTTTCATTATCTCCCAATAGC (TGAAAGCT)TTAGCGAAGCGATCMGCC GCC [X] CCCAATACITTTTATTGATTCACCATTTTC	28 27 29 28	80 74 60 74
54	FC1 RN1 FC2 RN2	21-48 352-327 345-376 1280-1255	GCC [X] CAATAAAACACAAAATGAATGAGTTAC (A)GATTTTGTTTTGAGCGTTAGAAATG (CAAAATC)TATAAATCAATCAAGICAAAATG GCC [X] GCATTTACCCCTAAAACTATAAAC	28 26 32 26	68 66 62 70
15	FC1 RN1 FC2 RN2	14-35 157-132 147-179 377-349	GCC [X] CTGAAGGGTGTATGGTATTAGG (C)ACCACATATGTATCTGCATTAATG (CATGTATGGT)GTAGCAAGAATTTTAAGGAGGC GCC [X] CGTTAAACIAAAGTCTATTTTAAATTC	22 26 33 29	64 68 64 70
57	FC1 RN1 FC2 RN2	13-39 267-244 258-294 957-934	GCC [X] GTAAGGAATGAGATGATAAAGAGTTGG (T)GGAATATCTGATCCAGCCATC (GAATATTCC)AAAAGCCGTTTTTTTATTACAGAAGAG GCC [X] CTAAGCTCTGGCTTATTGCGTATC	27 24 37 24	74 68 76 68
87	FC1 RN1 FC2 RN2	1-22 27-3 18-50 519-498	GCC [X] ATGCGTTTATTATTGGTGGG (C)AATACCCACCAATATAAACGCAT (GTGGGATTT)GGTATTATCGCTCTTTTAAATCC GCC [X] TTAAATTTTAGGGAAAGGGTA	22 25 33 22	62 66 64 62

CONDITIONS FOR RECOMBINANT PCR

Two independent PCR conditions are carried out for FC1/RN1 and FC2/RN2 primers under the same conditions proposed for cloning genes for expression.

After 20 cycles, the product of each reaction is used as template for a further 20 cycles with FC1/RN2 only. The product will encompass the full length gene minus the transposon. The presence of restriction sites at the 5' ends of these primers allows for cloning/expression studies.

-74-

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: ORAVAX, INC.

(ii) TITLE OF THE INVENTION: HELICOBACTER POLYPEPTIDES
AND CORRESPONDING POLYNUCLEOTIDE MOLECULES

(iii) NUMBER OF SEQUENCES: 190

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Clark & Elbing LLP
- (B) STREET: 176 Federal Street
- (C) CITY: Boston
- (D) STATE: MA
- (E) COUNTRY: USA
- (F) ZIP: 02110-2214

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: DOS
- (D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: UNKNOWN
- (B) FILING DATE: 14-NOV-1997
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/749,051
- (B) FILING DATE: 14-NOV-1996

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/831,309
- (B) FILING DATE: 1-APR-1997

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/834,705
- (B) FILING DATE: 1-APR-1997

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/833,457
- (B) FILING DATE: 1-APR-1997

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/881,227
- (B) FILING DATE: 24-JUN-1997

(vii) PRIOR APPLICATION DATA:

-75-

(A) APPLICATION NUMBER: 08/902,615

(B) FILING DATE: 29-JUL-1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Clark, Paul T.

(B) REGISTRATION NUMBER: 30,175

(C) REFERENCE/DOCKET NUMBER: 06132/028WO1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-428-0200

(B) TELEFAX: 617-428-7045

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 989 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 71...940

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTATGACGAT	TGTCTCGCTT	TTAGAAAACA	CTCTAATCGC	TTTTGAAAAA	CAACAAAGGA	60
AGGGATTTTA	ATG AAA TTT TTA	CGC TCT GTT TAT	GCA TTT TGC TCC	AGT		109
	Met Lys Phe Leu Arg	Ser Val Tyr Ala	Phe Cys Ser Ser			
	1	5	10			
TGG GTA GGG	ACG ATT GTT	ATT GTG CTG	TTG GTT ATC	TTT TTT ATC	GCG	157
Trp Val Gly	Thr Ile Val	Ile Val Leu	Leu Val Ile	Phe Phe Ile	Ala	
	15	20	25			
CAA GCC TTT	ATC ATT CCC	TCT CGC TCT	ATG GTT GGC	ACG CTC TAT	GAG	205
Gln Ala Phe	Ile Ile Pro	Ser Arg Ser	Met Val Gly	Thr Leu Tyr	Glu	
	30	35	40	45		
GGC GAC ATG	CTC TTT GTC	AAA AAG TTT	TCT TAC GGC	ATA CCC ATT	CCT	253
Gly Asp Met	Leu Phe Val	Lys Lys Phe	Ser Tyr Gly	Ile Pro Ile	Pro	
	50	55	60			
AAA ATC CCA	TGG ATT GAG	CTT CCT GTT	ATG CCT GAT	TTT AAA AAT	AAC	301
Lys Ile Pro	Trp Ile Glu	Leu Pro Val	Met Pro Asp	Phe Lys Asn	Asn	
	65	70	75			
GGA CAT TTG	ATA GAG GGG	GAT CGC CCT	AAG CGT GGC	GAA GTG GTG	GTG	349
Gly His Leu	Ile Glu Gly	Asp Arg Pro	Lys Arg Gly	Glu Val Val	Val	

-76-

80	85	90	
TTT ATC CCT CCC CAT GAA AAA AAG TCT TAC TAT GTT AAA AGG AAT TTT			397
Phe Ile Pro Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe			
95	100	105	
GCC ATT GGA GGC GAT GAG GTG TTG TTC ACT AAT GAG GGT TTT TAT TTG			445
Ala Ile Gly Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu			
110	115	120	125
CAC CCT TTT GAG AGC GAC ACG GAC AAA AAT TAC ATC GCT AAA CAT TAC			493
His Pro Phe Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr			
	130	135	140
CCT AAC GCC ATG ACA AAA GAA TTT ATG GGT AAA ATT TTT GTT TTA AAC			541
Pro Asn Ala Met Thr Lys Glu Phe Met Gly Lys Ile Phe Val Leu Asn			
	145	150	155
CCT TAT AAA AAT GAG CAT CCG GGT ATC CAT TAC CAA AAA GAC AAT GAA			589
Pro Tyr Lys Asn Glu His Pro Gly Ile His Tyr Gln Lys Asp Asn Glu			
	160	165	170
ACC TTC CAC TTA ATG GAG CAA TTA GCC ACT CAA GGC GCA GAA GCT AAT			637
Thr Phe His Leu Met Glu Gln Leu Ala Thr Gln Gly Ala Glu Ala Asn			
	175	180	185
ATC AGC ATG CAA CTC ATT CAA ATG GAG GGC GAA AAG GTG TTT TAT AAG			685
Ile Ser Met Gln Leu Ile Gln Met Glu Gly Glu Lys Val Phe Tyr Lys			
190	195	200	205
AAA ATC AAT GAC GAT GAA TTT TTC ATG ATC GGC GAC AAC AGA GAC AAT			733
Lys Ile Asn Asp Asp Glu Phe Phe Met Ile Gly Asp Asn Arg Asp Asn			
	210	215	220
TCT AGC GAC TCG CGC TTT TGG GGG AGT GTG GCT TAT AAA AAC ATC GTG			781
Ser Ser Asp Ser Arg Phe Trp Gly Ser Val Ala Tyr Lys Asn Ile Val			
	225	230	235
GGT TCG CCA TGG TTT GTT TAT TTC AGT TTG AGT TTA AAA AAT AGC CTA			829
Gly Ser Pro Trp Phe Val Tyr Phe Ser Leu Ser Leu Lys Asn Ser Leu			
	240	245	250
GAA ATG GAT GCA GAA AAT AAC CCT AAA AAA CGC TAT CTG GTG CGT TGG			877
Glu Met Asp Ala Glu Asn Asn Pro Lys Lys Arg Tyr Leu Val Arg Trp			
	255	260	265
GAA CGC ATG TTT AAA AGC GTT GGA GGC TTA GAA AAA ATC ATT AAA AAA			925
Glu Arg Met Phe Lys Ser Val Gly Gly Leu Glu Lys Ile Ile Lys Lys			
270	275	280	285
GAA AAC GCA ACG CAT TAAGGTTTTT TGTGCAATTT TTTGATTCT CTTTAGAAAG T			981
Glu Asn Ala Thr His			
	290		
TTTATTAC			989

-77-

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Phe Leu Arg Ser Val Tyr Ala Phe Cys Ser Ser Trp Val Gly
 1             5             10             15
Thr Ile Val Ile Val Leu Leu Val Ile Phe Phe Ile Ala Gln Ala Phe
      20             25             30
Ile Ile Pro Ser Arg Ser Met Val Gly Thr Leu Tyr Glu Gly Asp Met
      35             40             45
Leu Phe Val Lys Lys Phe Ser Tyr Gly Ile Pro Ile Pro Lys Ile Pro
      50             55             60
Trp Ile Glu Leu Pro Val Met Pro Asp Phe Lys Asn Asn Gly His Leu
      65             70             75             80
Ile Glu Gly Asp Arg Pro Lys Arg Gly Glu Val Val Val Phe Ile Pro
      85             90             95
Pro His Glu Lys Lys Ser Tyr Tyr Val Lys Arg Asn Phe Ala Ile Gly
      100            105            110
Gly Asp Glu Val Leu Phe Thr Asn Glu Gly Phe Tyr Leu His Pro Phe
      115            120            125
Glu Ser Asp Thr Asp Lys Asn Tyr Ile Ala Lys His Tyr Pro Asn Ala
      130            135            140
Met Thr Lys Glu Phe Met Gly Lys Ile Phe Val Leu Asn Pro Tyr Lys
      145            150            155            160
Asn Glu His Pro Gly Ile His Tyr Gln Lys Asp Asn Glu Thr Phe His
      165            170            175
Leu Met Glu Gln Leu Ala Thr Gln Gly Ala Glu Ala Asn Ile Ser Met
      180            185            190
Gln Leu Ile Gln Met Glu Gly Glu Lys Val Phe Tyr Lys Lys Ile Asn
      195            200            205
Asp Asp Glu Phe Phe Met Ile Gly Asp Asn Arg Asp Asn Ser Ser Asp
      210            215            220
Ser Arg Phe Trp Gly Ser Val Ala Tyr Lys Asn Ile Val Gly Ser Pro
      225            230            235            240
Trp Phe Val Tyr Phe Ser Leu Ser Leu Lys Asn Ser Leu Glu Met Asp
      245            250            255
Ala Glu Asn Asn Pro Lys Lys Arg Tyr Leu Val Arg Trp Glu Arg Met
      260            265            270
Phe Lys Ser Val Gly Gly Leu Glu Lys Ile Ile Lys Lys Glu Asn Ala
      275            280            285
Thr His
      290

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

-78-

- (A) LENGTH: 514 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 112...471
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

GGATTTT TTA GAGCTCTTAG TCAATGATAA TGTGGTAGAA ACGATTGAAA AAGGCTTTGT 60
GATAGGTTTT GGAGCGGGG ATATTACCTA TCAATTAAGA GGCGAAATGT A ATG GGT 117
                                         Met Gly
                                         1

GCA GTG GTT GTT TTA TTT TTA ACG CTG GTT TTA TTG TTT TTA GTT TTA 165
Ala Val Val Val Leu Phe Leu Thr Leu Val Leu Leu Phe Leu Val Leu
      5                      10                      15

AGG GAT TTT GGT TTA GCA AGC CCC AAA CAA AAG ATT TTA GCT TTT TTA 213
Arg Asp Phe Gly Leu Ala Ser Pro Lys Gln Lys Ile Leu Ala Phe Leu
      20                      25                      30

ATC GTA GGG ATT ATA GGA GCG AGC ATC AGC GTT TAT ACT TAC AAG CAA 261
Ile Val Gly Ile Ile Gly Ala Ser Ile Ser Val Tyr Thr Tyr Lys Gln
      35                      40                      45                      50

AAC CAA CAA AAC CAA CAA GAG ATC GCT TTG CAA AGA GCG TTT TTA AGG 309
Asn Gln Gln Asn Gln Gln Glu Ile Ala Leu Gln Arg Ala Phe Leu Arg
      55                      60                      65

GGG GAA ACC TTG TTG TGT AAA GGC ATT AAA GTC AAT AAC CAA ACC TTT 357
Gly Glu Thr Leu Leu Cys Lys Gly Ile Lys Val Asn Asn Gln Thr Phe
      70                      75                      80

AAT TTA GTG AGC GGA ACT TTA AGC TTT TTA GGC AAA AAA CAA ACC CCT 405
Asn Leu Val Ser Gly Thr Leu Ser Phe Leu Gly Lys Lys Gln Thr Pro
      85                      90                      95

ATG AAA GAC GTT CTT GTG GAT TTG GAT TCT TGT CAG ACG CTC CAA AAA 453
Met Lys Asp Val Leu Val Asp Leu Asp Ser Cys Gln Thr Leu Gln Lys
      100                      105                      110

GAT CCC TTA ATC CAA CCC TAATGATGAA TAATAATAAT ACCCCACCCA AACCCCTA 509
Asp Pro Leu Ile Gln Pro
      115                      120

GAAGA 514

```

(2) INFORMATION FOR SEQ ID NO:4:

-79-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Gly Ala Val Val Val Leu Phe Leu Thr Leu Val Leu Leu Phe Leu
 1               5               10               15
Val Leu Arg Asp Phe Gly Leu Ala Ser Pro Lys Gln Lys Ile Leu Ala
 20               25               30
Phe Leu Ile Val Gly Ile Ile Gly Ala Ser Ile Ser Val Tyr Thr Tyr
 35               40               45
Lys Gln Asn Gln Gln Asn Gln Gln Glu Ile Ala Leu Gln Arg Ala Phe
 50               55               60
Leu Arg Gly Glu Thr Leu Leu Cys Lys Gly Ile Lys Val Asn Asn Gln
 65               70               75               80
Thr Phe Asn Leu Val Ser Gly Thr Leu Ser Phe Leu Gly Lys Lys Gln
 85               90               95
Thr Pro Met Lys Asp Val Leu Val Asp Leu Asp Ser Cys Gln Thr Leu
 100              105              110
Gln Lys Asp Pro Leu Ile Gln Pro
 115              120

```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 135...1049
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

GTTTTTAATT TAATATTCAT TAAGCTTTTG TGGCTATTCC ATTTTAATTT TGTTTTTCAT      60
TAAACCCCAA TCTAAATCT TATTTTATG ATAAATACC TAATCATAAT ATCAAATCTT      120
AAACCAACGA AACC ATG AAA AAA GCT CTC TTA CTA ACT CTC TCT CTC TCG      170
          Met Lys Lys Ala Leu Leu Leu Thr Leu Ser Leu Ser
              1               5               10

TTC TGG CTC CAC GCT GAA AGG AAT GGA TTT TAT TTA GGT TTA AAT TTT      218
Phe Trp Leu His Ala Glu Arg Asn Gly Phe Tyr Leu Gly Leu Asn Phe
 15               20               25

```


-80-

CTA GAA GGA AGC TAT ATT AAA GGA CAA GGT AGC ATC GGC AAA AAA GCT Leu Glu Gly Ser Tyr Ile Lys Gly Gln Gly Ser Ile Gly Lys Lys Ala 30 35 40	266
TCA GCA GAA AAC GCC TTA AAT GAA GCG ATC AAT AAC GCA AAA AAT TCA Ser Ala Glu Asn Ala Leu Asn Glu Ala Ile Asn Asn Ala Lys Asn Ser 45 50 55 60	314
TTA TTC CCT AAC ACA AAA GCC ATA AGA GAT GCA CAA AAC GCC TTA AAT Leu Phe Pro Asn Thr Lys Ala Ile Arg Asp Ala Gln Asn Ala Leu Asn 65 70 75	362
GCA GTG AAA GAT TCA AAC AAA ATC GCT AGC CGA TTC GCA GGA AAT GGT Ala Val Lys Asp Ser Asn Lys Ile Ala Ser Arg Phe Ala Gly Asn Gly 80 85 90	410
GGA TCG GGC GGT CTT TTT AAT GAG CTC AGC TTT GGG TAT AAA TAT TTT Gly Ser Gly Gly Leu Phe Asn Glu Leu Ser Phe Gly Tyr Lys Tyr Phe 95 100 105	458
TTG GGT AAA AAA AGG ATT ATA GGG TTT AGG CAC TCT CTT TTT TTC GGT Leu Gly Lys Lys Arg Ile Ile Gly Phe Arg His Ser Leu Phe Phe Gly 110 115 120	506
TAC CAA CTT GGT GGC GTT GGT TCT GTT CCT GGT AGC GGT TTA ATC GTT Tyr Gln Leu Gly Gly Val Gly Ser Val Pro Gly Ser Gly Leu Ile Val 125 130 135 140	554
TTT TTA CCC TAT GGT TTC AAT ACG GAT TTG CTC ATT AAT TGG ACT AAC Phe Leu Pro Tyr Gly Phe Asn Thr Asp Leu Leu Ile Asn Trp Thr Asn 145 150 155	602
GAT AAG CGA GCG TCC CAA AAA TAT GTT GAA CGA AGG GTA AAA GGG CTC Asp Lys Arg Ala Ser Gln Lys Tyr Val Glu Arg Arg Val Lys Gly Leu 160 165 170	650
TCT ATA TTT TAC AAA GAT ATG ACC GGC AGA ACG CTA GAC GCT AAT ACA Ser Ile Phe Tyr Lys Asp Met Thr Gly Arg Thr Leu Asp Ala Asn Thr 175 180 185	698
TTA AAA AAA GCA TCA AGG CAT GTA TTT AGA AAA TCT TCA GGG CTT GTG Leu Lys Lys Ala Ser Arg His Val Phe Arg Lys Ser Ser Gly Leu Val 190 195 200	746
ATT GGC ATG GAA CTA GGG GGT AGC ACT TGG TTT GCA AGT AAC AAT CTC Ile Gly Met Glu Leu Gly Gly Ser Thr Trp Phe Ala Ser Asn Asn Leu 205 210 215 220	794
ACC CCT TTC AAT CAA GTC AAG AGT CGC ACG ATT TTT CAG TTG CAA GGA Thr Pro Phe Asn Gln Val Lys Ser Arg Thr Ile Phe Gln Leu Gln Gly 225 230 235	842
AAA TTT GGC GTT CGT TGG AAT AAT GAT GAA TAC GAT ATT GAT CGC TAT Lys Phe Gly Val Arg Trp Asn Asn Asp Glu Tyr Asp Ile Asp Arg Tyr 240 245 250	890

-81-

```

GGC GAT GAA ATC TAT CTT GGA GGT TCT AGT GTT GAA TTA GGG GTT AAA      938
Gly Asp Glu Ile Tyr Leu Gly Gly Ser Ser Val Glu Leu Gly Val Lys
      255                      260                      265

GTG CCA GCG TTT AAA GTC AAT TAC TAT AGC GAT GAT TAT GGG GAT AAA      986
Val Pro Ala Phe Lys Val Asn Tyr Tyr Ser Asp Asp Tyr Gly Asp Lys
      270                      275                      280

TTG GAT TAT AAA AGA GTG GTG AGC GTT TAT CTT AAC TAT ACA TAT AAC      1034
Leu Asp Tyr Lys Arg Val Val Ser Val Tyr Leu Asn Tyr Thr Tyr Asn
      285                      290                      295                      300

TTT AAA AAC AAA CAT TAAAACACGC TTTTACC GC TCTTTAGTTG GTTTTTTAAA A      1090
Phe Lys Asn Lys His
                        305

AACCTTATTT TTTATTAGCT TGAAACTCTT CAAAGCCTTT TTTTCTCAAT TGGCATGCCG      1150
GGCATTATATC GCAACCATAA CCATAAGCAT GCAAAATCTT TCGCTCTCCT TGATAGCAGG      1210
TGTGCGTTTC TTTGATGACT AAA                                           1233

```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Lys Lys Ala Leu Leu Leu Thr Leu Ser Leu Ser Phe Trp Leu His
 1          5          10          15
Ala Glu Arg Asn Gly Phe Tyr Leu Gly Leu Asn Phe Leu Glu Gly Ser
      20          25          30
Tyr Ile Lys Gly Gln Gly Ser Ile Gly Lys Lys Ala Ser Ala Glu Asn
      35          40          45
Ala Leu Asn Glu Ala Ile Asn Asn Ala Lys Asn Ser Leu Phe Pro Asn
      50          55          60
Thr Lys Ala Ile Arg Asp Ala Gln Asn Ala Leu Asn Ala Val Lys Asp
      65          70          75          80
Ser Asn Lys Ile Ala Ser Arg Phe Ala Gly Asn Gly Gly Ser Gly Gly
      85          90          95
Leu Phe Asn Glu Leu Ser Phe Gly Tyr Lys Tyr Phe Leu Gly Lys Lys
      100         105         110
Arg Ile Ile Gly Phe Arg His Ser Leu Phe Phe Gly Tyr Gln Leu Gly
      115         120         125
Gly Val Gly Ser Val Pro Gly Ser Gly Leu Ile Val Phe Leu Pro Tyr
      130         135         140
Gly Phe Asn Thr Asp Leu Leu Ile Asn Trp Thr Asn Asp Lys Arg Ala
      145         150         155         160
Ser Gln Lys Tyr Val Glu Arg Arg Val Lys Gly Leu Ser Ile Phe Tyr
      165         170         175

```

-82-

Lys Asp Met Thr Gly Arg Thr Leu Asp Ala Asn Thr Leu Lys Lys Ala
 180 185 190
 Ser Arg His Val Phe Arg Lys Ser Ser Gly Leu Val Ile Gly Met Glu
 195 200 205
 Leu Gly Gly Ser Thr Trp Phe Ala Ser Asn Asn Leu Thr Pro Phe Asn
 210 215 220
 Gln Val Lys Ser Arg Thr Ile Phe Gln Leu Gln Gly Lys Phe Gly Val
 225 230 235 240
 Arg Trp Asn Asn Asp Glu Tyr Asp Ile Asp Arg Tyr Gly Asp Glu Ile
 245 250 255
 Tyr Leu Gly Gly Ser Ser Val Glu Leu Gly Val Lys Val Pro Ala Phe
 260 265 270
 Lys Val Asn Tyr Tyr Ser Asp Asp Tyr Gly Asp Lys Leu Asp Tyr Lys
 275 280 285
 Arg Val Val Ser Val Tyr Leu Asn Tyr Thr Tyr Asn Phe Lys Asn Lys
 290 295 300
 His
 305

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3012 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 142...2682
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATGACGGCT	CTAAACCAAA	CGATTTGACT	TCTCCAAAAG	AAGCCTCTCA	AGAATCTCAA	60
AAAAATGAAG	CTCCAAAAAA	TGAAGTTCAA	AGAAATGAAG	CTCAAAAAGA	AACCCCCCAA	120
TCCAATCAAA	CGCCTAAAGA	A ATG AAA	GTC AAG TCC	ATT TCT TAT	GTC GGG	171
	Met Lys Val	Lys Ser	Ile Ser Tyr	Val Gly		
	1	5		10		
CTT TCT TAC	ATG TCT GAC	ATG CTC GCT	AAT GAA ATT	GTA AAG ATT	CGT	219
Leu Ser Tyr	Met Ser Asp	Met Leu Ala	Asn Glu Ile	Val Lys Ile	Arg	
	15	20		25		
GTG GGC GAT	ATT GTG GAT	TCT AAA AAA	ATA GAC ACC	GCT GTT TTG	GCT	267
Val Gly Asp	Ile Val Asp	Ser Lys Lys	Ile Asp Thr	Ala Val Leu	Ala	
	30	35		40		
TTG TTC AAT	CAA GGG TAT	TTT AAA GAC	GTT TAT GCC	ACT TTT GAA	GGC	315
Leu Phe Asn	Gln Gly Tyr	Phe Lys Asp	Val Tyr Ala	Thr Phe Glu	Gly	
	45	50		55		

- 83 -

GGC ATA TTA GAG TTT CAT TTT GAT GAA AAA GCC AGG ATT GCC GGG GTA	363
Gly Ile Leu Glu Phe His Phe Asp Glu Lys Ala Arg Ile Ala Gly Val	
60 65 70	
GAA ATC AAG GGT TAT GGG ACT GAA AAG GAA AAA GAC GGC TTA AAA TCC	411
Glu Ile Lys Gly Tyr Gly Thr Glu Lys Glu Lys Asp Gly Leu Lys Ser	
75 80 85 90	
CAA ATG GGG ATC AAA AAG GGC GAC ACC TTT GAT GAG CAA AAA TTA GAG	459
Gln Met Gly Ile Lys Lys Gly Asp Thr Phe Asp Glu Gln Lys Leu Glu	
95 100 105	
CAT GCT AAA ACG GCT TTA AAA ACC GCT TTA GAG GGG CAG GGC TAT TAT	507
His Ala Lys Thr Ala Leu Lys Thr Ala Leu Glu Gly Gln Gly Tyr Tyr	
110 115 120	
GGG AGC GTG GTG GAG GTG CGC ACA GAA AAG GTC AGT GAG GGT GCA TTA	555
Gly Ser Val Val Glu Val Arg Thr Glu Lys Val Ser Glu Gly Ala Leu	
125 130 135	
TTG ATC GTG TTT GAT GTG AAT AGG GGG GAT AGC ATT TAT ATC AAA CAA	603
Leu Ile Val Phe Asp Val Asn Arg Gly Asp Ser Ile Tyr Ile Lys Gln	
140 145 150	
TCC ATT TAT GAG GGA AGC GCG AAA TTA AAA CGC CGC ATG ATT GAA TCT	651
Ser Ile Tyr Glu Gly Ser Ala Lys Leu Lys Arg Arg Met Ile Glu Ser	
155 160 165 170	
TTG AGT GCG AAC AAG CAA CGA GAT TTC ATG GGC TGG ATG TGG GGC TTG	699
Leu Ser Ala Asn Lys Gln Arg Asp Phe Met Gly Trp Met Trp Gly Leu	
175 180 185	
AAT GAC GGG AAA TTG CGT TTA GAT CAA CTA GAA TAC GAT TCT ATG CGT	747
Asn Asp Gly Lys Leu Arg Leu Asp Gln Leu Glu Tyr Asp Ser Met Arg	
190 195 200	
ATC CAA GAT GTG TAT ATG CGT AGG GGT TAC TTA GAC GCT CAT ATT TCT	795
Ile Gln Asp Val Tyr Met Arg Arg Gly Tyr Leu Asp Ala His Ile Ser	
205 210 215	
TCG CCT TTT TTG AAA ACG GAT TTT TCT ACC CAT GAC GCT AAG CTT CAT	843
Ser Pro Phe Leu Lys Thr Asp Phe Ser Thr His Asp Ala Lys Leu His	
220 225 230	
TAT AAA GTC AAA GAG GGG ATC CAA TAC AGG ATT TCA GAC ATT TTA ATA	891
Tyr Lys Val Lys Glu Gly Ile Gln Tyr Arg Ile Ser Asp Ile Leu Ile	
235 240 245 250	
GAG ATT GAC AAC CCG GTA GTC CCC TTA AAA ACC TTA GAA AAA GCG CTT	939
Glu Ile Asp Asn Pro Val Val Pro Leu Lys Thr Leu Glu Lys Ala Leu	
255 260 265	
AAA GTG AAA AGG AAA GAT GTC TTT AAT ATT GAG CAT TTA AGA GCG GAT	987
Lys Val Lys Arg Lys Asp Val Phe Asn Ile Glu His Leu Arg Ala Asp	
270 275 280	

- 84 -

GCG CAA ATT TTA AAA ACC GAA ATC GCC GAT AAG GGT TAT GCG TTT GCG	1035
Ala Gln Ile Leu Lys Thr Glu Ile Ala Asp Lys Gly Tyr Ala Phe Ala	
285 290 295	
GTG GTG AAG CCA GAC TTG GAT AAA GAT GAA AAA AAC GGG CTT GTG AAA	1083
Val Val Lys Pro Asp Leu Asp Lys Asp Glu Lys Asn Gly Leu Val Lys	
300 305 310	
GTC ATT TAT CGT ATT GAA GTG GGC GAT ATG GTG TAT ATC AAT GAT GTC	1131
Val Ile Tyr Arg Ile Glu Val Gly Asp Met Val Tyr Ile Asn Asp Val	
315 320 325 330	
ATC ATT TCA GGG AAC CAG CGC ACG AGC GAT AGG ATC ATT AGA AGG GAG	1179
Ile Ile Ser Gly Asn Gln Arg Thr Ser Asp Arg Ile Ile Arg Arg Glu	
335 340 345	
TTA TTG TTA GGG CCT AAG GAT AAA TAC AAC TTG ACC AAA CTG AGA AAT	1227
Leu Leu Leu Gly Pro Lys Asp Lys Tyr Asn Leu Thr Lys Leu Arg Asn	
350 355 360	
TCC GAA AAT TCT TTA AGG CGT TTA GGA TTC TTC TCT AAA GTC AAA ATT	1275
Ser Glu Asn Ser Leu Arg Arg Leu Gly Phe Phe Ser Lys Val Lys Ile	
365 370 375	
GAA GAA AAA AGG GTT AAT AGC TCA CTC ATG GAT TTA TTA GTG AGC GTA	1323
Glu Glu Lys Arg Val Asn Ser Ser Leu Met Asp Leu Leu Val Ser Val	
380 385 390	
GAA GAG GGG CGT ACT GGG CAG TTG CAA TTT GGG TTA GGC TAT GGC TCT	1371
Glu Glu Gly Arg Thr Gly Gln Leu Gln Phe Gly Leu Gly Tyr Gly Ser	
395 400 405 410	
TAT GGA GGG CTT ATG CTT AAT GGG AGC GTG AGC GAA AGA AAC CTT TTT	1419
Tyr Gly Gly Leu Met Leu Asn Gly Ser Val Ser Glu Arg Asn Leu Phe	
415 420 425	
GGC ACA GGG CAA AGC ATG AGC TTG TAT GCT AAC ATC GCT ACA GGG GGG	1467
Gly Thr Gly Gln Ser Met Ser Leu Tyr Ala Asn Ile Ala Thr Gly Gly	
430 435 440	
GGT AGA TCT TAT CCG GGC ATG CCA AAA GGA GCG GGG CGT ATG TTT GCC	1515
Gly Arg Ser Tyr Pro Gly Met Pro Lys Gly Ala Gly Arg Met Phe Ala	
445 450 455	
GGG AAT TTG AGC TTG ACT AAT CCA AGG ATT TTT GAC AGC TGG TAT AGC	1563
Gly Asn Leu Ser Leu Thr Asn Pro Arg Ile Phe Asp Ser Trp Tyr Ser	
460 465 470	
TCT ACG ATC AAC CTT TAT GCG GAT TAC AGG ATA AGC TAC CAA TAC ATC	1611
Ser Thr Ile Asn Leu Tyr Ala Asp Tyr Arg Ile Ser Tyr Gln Tyr Ile	
475 480 485 490	
CAA CAA GGC GGG GGC TTT GGG GTG AAT GTC GGG CGC ATG CTG GGT AAT	1659
Gln Gln Gly Gly Gly Phe Gly Val Asn Val Gly Arg Met Leu Gly Asn	
495 500 505	

- 85 -

AGA ACC CAT GTG AGC TTA GGG TAT AAC TTG AAT GTT ACC AAA CTC CTT	1707
Arg Thr His Val Ser Leu Gly Tyr Asn Leu Asn Val Thr Lys Leu Leu	
510 515 520	
GGT TTC AGC AGC CCT TTA TAC AAC CGC TAC TAT TCC TCT GTT AAT GAA	1755
Gly Phe Ser Ser Pro Leu Tyr Asn Arg Tyr Tyr Ser Ser Val Asn Glu	
525 530 535	
GTG GTT TCT CCA AGG CAA TGT TCT ACC CCC GCA TCG GTG ATT ATC AAT	1803
Val Val Ser Pro Arg Gln Cys Ser Thr Pro Ala Ser Val Ile Ile Asn	
540 545 550	
CGC TTA TCA GGC GGT AAA ACC CCC TTA CAA CCT GAA AGC TGT TCT AGT	1851
Arg Leu Ser Gly Gly Lys Thr Pro Leu Gln Pro Glu Ser Cys Ser Ser	
555 560 565 570	
CCT GGA GCG ATC ACC ACT TCA CCA GAA ATA AGA GGT ATT TGG GAT AGG	1899
Pro Gly Ala Ile Thr Thr Ser Pro Glu Ile Arg Gly Ile Trp Asp Arg	
575 580 585	
GAT TAC CAT ACG CCT ATC ACC AGC TCT TTC ACC CTT GAT GTG AGC TAT	1947
Asp Tyr His Thr Pro Ile Thr Ser Ser Phe Thr Leu Asp Val Ser Tyr	
590 595 600	
GAC AAC ACC GAT GAT TAT TAC TTC CCT AGA AAT GGG GTT ATC TTT AGT	1995
Asp Asn Thr Asp Asp Tyr Tyr Phe Pro Arg Asn Gly Val Ile Phe Ser	
605 610 615	
TCC TAT GCG ACG ATG TCT GGC TTG CCA AGC TCT GGC ACG CTC AAT TCT	2043
Ser Tyr Ala Thr Met Ser Gly Leu Pro Ser Ser Gly Thr Leu Asn Ser	
620 625 630	
TGG AAC GGG TTA GGC GGG AAT GTC CGT AAC ACC AAA GTT TAT GGT AAA	2091
Trp Asn Gly Leu Gly Gly Asn Val Arg Asn Thr Lys Val Tyr Gly Lys	
635 640 645 650	
TTC GCC GCT TAC CAC CAT TTG CAA AAA TAT TTA TTG ATA GAT TTG ATC	2139
Phe Ala Ala Tyr His His Leu Gln Lys Tyr Leu Leu Ile Asp Leu Ile	
655 660 665	
GCT CGC TTT AAA ACG CAA GGA GGT TAT ATC TTT AGG TAT AAC ACC GAT	2187
Ala Arg Phe Lys Thr Gln Gly Gly Tyr Ile Phe Arg Tyr Asn Thr Asp	
670 675 680	
GAT TAC TTG CCC TTA AAC TCC ACC TTC TAC ATG GGG GGC GTA ACC ACG	2235
Asp Tyr Leu Pro Leu Asn Ser Thr Phe Tyr Met Gly Gly Val Thr Thr	
685 690 695	
GTG AGA GGC TTT AGG AAC GGA TCG GTT ACT CCT AAA GAT GAG TTT GGC	2283
Val Arg Gly Phe Arg Asn Gly Ser Val Thr Pro Lys Asp Glu Phe Gly	
700 705 710	
TTG TGG CTT GGA GGC GAT GGG ATT TTT ACC GCT TCT ACT GAA TTG AGC	2331
Leu Trp Leu Gly Gly Asp Gly Ile Phe Thr Ala Ser Thr Glu Leu Ser	
715 720 725 730	

- 86 -

TAT GGG GTG CTA AAG GCG GCT AAA ATG CGC TTA GCG TGG TTT TTT GAC	2379
Tyr Gly Val Leu Lys Ala Ala Lys Met Arg Leu Ala Trp Phe Phe Asp	
735 740 745	
TTT GGT TTC TTA ACC TTT AAA ACC CCA ACT AGA GGG AGT TTT TTC TAT	2427
Phe Gly Phe Leu Thr Phe Lys Thr Pro Thr Arg Gly Ser Phe Phe Tyr	
750 755 760	
AAC GCT CCT GTT ACG ACA GCG AAT TTT AAA GAT TAT GGC GTT ATA GGG	2475
Asn Ala Pro Val Thr Thr Ala Asn Phe Lys Asp Tyr Gly Val Ile Gly	
765 770 775	
GCT GGG TTT GAA AGA GCG ACT TGG AGG GCT TCC ACA GGC TTG CAG ATT	2523
Ala Gly Phe Glu Arg Ala Thr Trp Arg Ala Ser Thr Gly Leu Gln Ile	
780 785 790	
GAA TGG ATT TCG CCC ATG GGG CCT TTG GTG TTG ATT TTC CCT ATA GCG	2571
Glu Trp Ile Ser Pro Met Gly Pro Leu Val Leu Ile Phe Pro Ile Ala	
795 800 805 810	
TTT TTC AAC CAA TGG GGC GAT GGC AAT GGC AAG AAA TGT AAA GGG CTA	2619
Phe Phe Asn Gln Trp Gly Asp Gly Asn Gly Lys Lys Cys Lys Gly Leu	
815 820 825	
TGC TTC AAC CCT AAC ATG GAC GAT TAC ACG CAA CAC TTT GAA TTT TCT	2667
Cys Phe Asn Pro Asn Met Asp Asp Tyr Thr Gln His Phe Glu Phe Ser	
830 835 840	
ATG GGA ACA AGG TTT TAAAATGCGC ATCAACAGAG AAGAAATTTT GGATTTAATG A	2723
Met Gly Thr Arg Phe	
845	
AAAACGCGCC CTTGAAAGAA TTGGGGCAAA GGGCTTTGAG GGTGAAGCAA CGCTTGACCC	2783
CTGAAAACCTT GACGACTTTT ATTGTGGATA GGAATATCAA TTACACCAAT ATTTGTTTTG	2843
TGGATTGCAA GTTTTGCGCG TTCAAACGCA CCTTAAAAGA AAAAGACGCC TATGTGTTGA	2903
GCTATGAAGA AATTGATCAA AAGATTGAAG AATTGCTCGC TATTGGCGGC ACGCAGATCC	2963
TTTTTCAAGG GGGGTGCAC CCGCAGCTAA AGATTGATTA TTATGAGAA	3012

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 847 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Val Lys Ser Ile Ser Tyr Val Gly Leu Ser Tyr Met Ser Asp	
1 5 10 15	
Met Leu Ala Asn Glu Ile Val Lys Ile Arg Val Gly Asp Ile Val Asp	
20 25 30	

-87-

Ser Lys Lys Ile Asp Thr Ala Val Leu Ala Leu Phe Asn Gln Gly Tyr
 35 40 45
 Phe Lys Asp Val Tyr Ala Thr Phe Glu Gly Gly Ile Leu Glu Phe His
 50 55 60
 Phe Asp Glu Lys Ala Arg Ile Ala Gly Val Glu Ile Lys Gly Tyr Gly
 65 70 75 80
 Thr Glu Lys Glu Lys Asp Gly Leu Lys Ser Gln Met Gly Ile Lys Lys
 85 90 95
 Gly Asp Thr Phe Asp Glu Gln Lys Leu Glu His Ala Lys Thr Ala Leu
 100 105 110
 Lys Thr Ala Leu Glu Gly Gln Gly Tyr Tyr Gly Ser Val Val Glu Val
 115 120 125
 Arg Thr Glu Lys Val Ser Glu Gly Ala Leu Leu Ile Val Phe Asp Val
 130 135 140
 Asn Arg Gly Asp Ser Ile Tyr Ile Lys Gln Ser Ile Tyr Glu Gly Ser
 145 150 155 160
 Ala Lys Leu Lys Arg Arg Met Ile Glu Ser Leu Ser Ala Asn Lys Gln
 165 170 175
 Arg Asp Phe Met Gly Trp Met Trp Gly Leu Asn Asp Gly Lys Leu Arg
 180 185 190
 Leu Asp Gln Leu Glu Tyr Asp Ser Met Arg Ile Gln Asp Val Tyr Met
 195 200 205
 Arg Arg Gly Tyr Leu Asp Ala His Ile Ser Ser Pro Phe Leu Lys Thr
 210 215 220
 Asp Phe Ser Thr His Asp Ala Lys Leu His Tyr Lys Val Lys Glu Gly
 225 230 235 240
 Ile Gln Tyr Arg Ile Ser Asp Ile Leu Ile Glu Ile Asp Asn Pro Val
 245 250 255
 Val Pro Leu Lys Thr Leu Glu Lys Ala Leu Lys Val Lys Arg Lys Asp
 260 265 270
 Val Phe Asn Ile Glu His Leu Arg Ala Asp Ala Gln Ile Leu Lys Thr
 275 280 285
 Glu Ile Ala Asp Lys Gly Tyr Ala Phe Ala Val Val Lys Pro Asp Leu
 290 295 300
 Asp Lys Asp Glu Lys Asn Gly Leu Val Lys Val Ile Tyr Arg Ile Glu
 305 310 315 320
 Val Gly Asp Met Val Tyr Ile Asn Asp Val Ile Ile Ser Gly Asn Gln
 325 330 335
 Arg Thr Ser Asp Arg Ile Ile Arg Arg Glu Leu Leu Leu Gly Pro Lys
 340 345 350
 Asp Lys Tyr Asn Leu Thr Lys Leu Arg Asn Ser Glu Asn Ser Leu Arg
 355 360 365
 Arg Leu Gly Phe Phe Ser Lys Val Lys Ile Glu Glu Lys Arg Val Asn
 370 375 380
 Ser Ser Leu Met Asp Leu Leu Val Ser Val Glu Glu Gly Arg Thr Gly
 385 390 395 400
 Gln Leu Gln Phe Gly Leu Gly Tyr Gly Ser Tyr Gly Gly Leu Met Leu
 405 410 415
 Asn Gly Ser Val Ser Glu Arg Asn Leu Phe Gly Thr Gly Gln Ser Met
 420 425 430
 Ser Leu Tyr Ala Asn Ile Ala Thr Gly Gly Gly Arg Ser Tyr Pro Gly
 435 440 445
 Met Pro Lys Gly Ala Gly Arg Met Phe Ala Gly Asn Leu Ser Leu Thr
 450 455 460
 Asn Pro Arg Ile Phe Asp Ser Trp Tyr Ser Ser Thr Ile Asn Leu Tyr

-88-

465					470					475					480
Ala	Asp	Tyr	Arg	Ile	Ser	Tyr	Gln	Tyr	Ile	Gln	Gln	Gly	Gly	Gly	Phe
				485					490					495	
Gly	Val	Asn	Val	Gly	Arg	Met	Leu	Gly	Asn	Arg	Thr	His	Val	Ser	Leu
			500					505					510		
Gly	Tyr	Asn	Leu	Asn	Val	Thr	Lys	Leu	Leu	Gly	Phe	Ser	Ser	Pro	Leu
		515					520					525			
Tyr	Asn	Arg	Tyr	Tyr	Ser	Ser	Val	Asn	Glu	Val	Val	Ser	Pro	Arg	Gln
	530					535					540				
Cys	Ser	Thr	Pro	Ala	Ser	Val	Ile	Ile	Asn	Arg	Leu	Ser	Gly	Gly	Lys
545					550				555					560	
Thr	Pro	Leu	Gln	Pro	Glu	Ser	Cys	Ser	Ser	Pro	Gly	Ala	Ile	Thr	Thr
			565					570					575		
Ser	Pro	Glu	Ile	Arg	Gly	Ile	Trp	Asp	Arg	Asp	Tyr	His	Thr	Pro	Ile
			580					585				590			
Thr	Ser	Ser	Phe	Thr	Leu	Asp	Val	Ser	Tyr	Asp	Asn	Thr	Asp	Asp	Tyr
		595				600					605				
Tyr	Phe	Pro	Arg	Asn	Gly	Val	Ile	Phe	Ser	Ser	Tyr	Ala	Thr	Met	Ser
	610				615					620					
Gly	Leu	Pro	Ser	Ser	Gly	Thr	Leu	Asn	Ser	Trp	Asn	Gly	Leu	Gly	Gly
625					630					635				640	
Asn	Val	Arg	Asn	Thr	Lys	Val	Tyr	Gly	Lys	Phe	Ala	Ala	Tyr	His	His
			645					650					655		
Leu	Gln	Lys	Tyr	Leu	Leu	Ile	Asp	Leu	Ile	Ala	Arg	Phe	Lys	Thr	Gln
		660						665					670		
Gly	Gly	Tyr	Ile	Phe	Arg	Tyr	Asn	Thr	Asp	Asp	Tyr	Leu	Pro	Leu	Asn
		675					680					685			
Ser	Thr	Phe	Tyr	Met	Gly	Gly	Val	Thr	Thr	Val	Arg	Gly	Phe	Arg	Asn
	690				695					700					
Gly	Ser	Val	Thr	Pro	Lys	Asp	Glu	Phe	Gly	Leu	Trp	Leu	Gly	Gly	Asp
705					710				715					720	
Gly	Ile	Phe	Thr	Ala	Ser	Thr	Glu	Leu	Ser	Tyr	Gly	Val	Leu	Lys	Ala
				725					730					735	
Ala	Lys	Met	Arg	Leu	Ala	Trp	Phe	Phe	Asp	Phe	Gly	Phe	Leu	Thr	Phe
		740						745				750			
Lys	Thr	Pro	Thr	Arg	Gly	Ser	Phe	Phe	Tyr	Asn	Ala	Pro	Val	Thr	Thr
		755					760					765			
Ala	Asn	Phe	Lys	Asp	Tyr	Gly	Val	Ile	Gly	Ala	Gly	Phe	Glu	Arg	Ala
	770					775					780				
Thr	Trp	Arg	Ala	Ser	Thr	Gly	Leu	Gln	Ile	Glu	Trp	Ile	Ser	Pro	Met
785				790					795					800	
Gly	Pro	Leu	Val	Leu	Ile	Phe	Pro	Ile	Ala	Phe	Phe	Asn	Gln	Trp	Gly
			805						810				815		
Asp	Gly	Asn	Gly	Lys	Lys	Cys	Lys	Gly	Leu	Cys	Phe	Asn	Pro	Asn	Met
		820						825				830			
Asp	Asp	Tyr	Thr	Gln	His	Phe	Glu	Phe	Ser	Met	Gly	Thr	Arg	Phe	
		835					840					845			

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1032 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

-89-

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 149...913
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGTTTTGTG	TTGCAAAAAC	AAAACAGACC	AATAAAGGCA	TCACTTTTAA	AAGCGTTGTT	60
TAGGGGGGTT	TGGTTATTGG	TGTTTGATTA	GAATAGGGTT	GTTTTTAATT	TTCTTTTAAG	120
AGGAGTTTTT	ACTTTTTTAA	GGGTTTTT	ATG GAT ATT	TAT GCG TTA	TAT ATA	172
			Met Asp Ile Tyr	Ala Leu Tyr	Ile	
			1		5	
GCG ATA GGG	CTT TTT ACT	GGC ATT CTA	TCA GGG ATT	TTT GGC ATT	GGT	220
Ala Ile Gly	Leu Phe Thr	Gly Ile Leu	Ser Gly Ile	Phe Gly Ile	Gly	
10		15		20		
GGG GGG TTG	ATC ATT GTC	CCT ATC ATG	CTC GCA ACC	GGG CAT TCT	TTT	268
Gly Gly Leu	Ile Ile Val	Pro Ile Met	Leu Ala Thr	Gly His Ser	Phe	
25		30		35	40	
GAA GAA TCC	ATT GGG ATT	TCC ATT TTG	CAA ATG GCG	CTT TCA TCG	TTC	316
Glu Glu Ser	Ile Gly Ile	Ser Ile Leu	Gln Met Ala	Leu Ser Ser	Phe	
	45		50		55	
GTG GGC TCT	GTT TTG AAT	TTC AAA AAA	AAA TCG CTT	GAT TTT TCT	TTA	364
Val Gly Ser	Val Leu Asn	Phe Lys Lys	Lys Ser Leu	Asp Phe Ser	Leu	
	60		65		70	
GGC TTG TTG	ATA GGG GCA	GGG GGG CTG	ATA GGG GCG	AGT TTT AGC	GGA	412
Gly Leu Leu	Ile Gly Ala	Gly Gly Leu	Ile Gly Ala	Ser Phe Ser	Gly	
75		80		85		
TTT GTT TTA	AAA ATC GTT	TCC AGT AAA	ATT TTA ATG	GTT ATT TTC	GCG	460
Phe Val Leu	Lys Ile Val	Ser Ser Lys	Ile Leu Met	Val Ile Phe	Ala	
90		95		100		
CTT TTA GTC	GTG TAT TCT	ATG ATC CAA	TTT GTT TTG	AAA CCC AAA	AAA	508
Leu Leu Val	Val Tyr Ser	Met Ile Gln	Phe Val Leu	Lys Pro Lys	Lys	
105		110		115	120	
AAA GAT TTG	ATA GCG GAT	ACT AAA CGC	TAT CAT CTG	CAA GGT TTG	AAA	556
Lys Asp Leu	Ile Ala Asp	Thr Lys Arg	Tyr His Leu	Gln Gly Leu	Lys	
	125		130		135	
TTA TTT TTA	ATT GGC ACG	CTC ACA GGG	TTT TTT GCT	ATC ACT TTA	GGG	604
Leu Phe Leu	Ile Gly Thr	Leu Thr Gly	Phe Phe Ala	Ile Thr Leu	Gly	
	140		145		150	
ATT GGT GGG	GGG ATG CTC	ATG GTG CCT	TTG ATG CAT	TAT TTT TTA	GGG	652
Ile Gly Gly	Gly Met Leu	Met Val Pro	Leu Met His	Tyr Phe Leu	Gly	
155		160		165		

-90-

TAT GAT TCT AAA AAA TGC GTG GCT CTA GGG TTA TTT TTC ATC TTG TTT	700
Tyr Asp Ser Lys Lys Cys Val Ala Leu Gly Leu Phe Phe Ile Leu Phe	
170 175 180	
TCT TCT ATT TCA GGA GCT TTT TCT TTA ATG TAT CAC CAC ATC ATC AAT	748
Ser Ser Ile Ser Gly Ala Phe Ser Leu Met Tyr His His Ile Ile Asn	
185 190 195 200	
AAA GAA GTG CTC TTA GCA GGG GCG ATT GTG GGA TTA GGA TCT GTT ATG	796
Lys Glu Val Leu Leu Ala Gly Ala Ile Val Gly Leu Gly Ser Val Met	
205 210 215	
GGC GTG AGC ATT GGG ATT AAA TGG ATC ATG GGG CTT TTG AAT GAA AAA	844
Gly Val Ser Ile Gly Ile Lys Trp Ile Met Gly Leu Leu Asn Glu Lys	
220 225 230	
ATG CAT AAA GCT TTG ATT TTA GGG GTG TAT GGT TTG TCG CTA TTG ATT	892
Met His Lys Ala Leu Ile Leu Gly Val Tyr Gly Leu Ser Leu Leu Ile	
235 240 245	
GTT TTA TAC AAA CTC TTT TTT TAATTGATGG TTTTATACCA CTACTATTTT AAGA	947
Val Leu Tyr Lys Leu Phe Phe	
250 255	
CCCCTAAGAG TTTCCCTTTA GAGTATTTGC ATTTGTGCGC TAATGAGAGC CATTTATTGA	1007
GATTGGATTT TGATGCGGCC AATTT	1032

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 255 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Asp Ile Tyr Ala Leu Tyr Ile Ala Ile Gly Leu Phe Thr Gly Ile	
1 5 10 15	
Leu Ser Gly Ile Phe Gly Ile Gly Gly Gly Leu Ile Ile Val Pro Ile	
20 25 30	
Met Leu Ala Thr Gly His Ser Phe Glu Glu Ser Ile Gly Ile Ser Ile	
35 40 45	
Leu Gln Met Ala Leu Ser Ser Phe Val Gly Ser Val Leu Asn Phe Lys	
50 55 60	
Lys Lys Ser Leu Asp Phe Ser Leu Gly Leu Leu Ile Gly Ala Gly Gly	
65 70 75 80	
Leu Ile Gly Ala Ser Phe Ser Gly Phe Val Leu Lys Ile Val Ser Ser	
85 90 95	
Lys Ile Leu Met Val Ile Phe Ala Leu Leu Val Val Tyr Ser Met Ile	
100 105 110	
Gln Phe Val Leu Lys Pro Lys Lys Lys Asp Leu Ile Ala Asp Thr Lys	

		115					120					125						
Arg	Tyr	His	Leu	Gln	Gly	Leu	Lys	Leu	Phe	Leu	Ile	Gly	Thr	Leu	Thr			
	130					135					140							
Gly	Phe	Phe	Ala	Ile	Thr	Leu	Gly	Ile	Gly	Gly	Gly	Met	Leu	Met	Val			
145					150					155					160			
Pro	Leu	Met	His	Tyr	Phe	Leu	Gly	Tyr	Asp	Ser	Lys	Lys	Cys	Val	Ala			
				165					170					175				
Leu	Gly	Leu	Phe	Phe	Ile	Leu	Phe	Ser	Ser	Ile	Ser	Gly	Ala	Phe	Ser			
			180					185					190					
Leu	Met	Tyr	His	His	Ile	Ile	Asn	Lys	Glu	Val	Leu	Leu	Ala	Gly	Ala			
	195						200					205						
Ile	Val	Gly	Leu	Gly	Ser	Val	Met	Gly	Val	Ser	Ile	Gly	Ile	Lys	Trp			
	210						215				220							
Ile	Met	Gly	Leu	Leu	Asn	Glu	Lys	Met	His	Lys	Ala	Leu	Ile	Leu	Gly			
225					230					235					240			
Val	Tyr	Gly	Leu	Ser	Leu	Leu	Ile	Val	Leu	Tyr	Lys	Leu	Phe	Phe				
				245					250					255				

(A) LENGTH: 1057 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 66...980
(D) OTHER INFORMATION:

AAGAGCATGC	GAGAGAGCAT	AGAGGAATTT	TTTAATCAAG	AAATGTTGCA	AAGTGAAGTG	60										
CCGTT	ATG	GGT	AGA	ATT	GAA	TCA	AAA	AAG	CGT	TTG	AAA	GCG	CTT	GTT	TTT	110
	Met	Gly	Arg	Ile	Glu	Ser	Lys	Lys	Arg	Leu	Lys	Ala	Leu	Val	Phe	
	1					5				10					15	
TTA	GCC	AGC	TTG	GGG	GTT	TTG	TGG	GGC	AAT	AGC	GCT	GAA	AAA	ACG	CCT	158
Leu	Ala	Ser	Leu	Gly	Val	Leu	Trp	Gly	Asn	Ser	Ala	Glu	Lys	Thr	Pro	
				20					25					30		
TTT	TTT	AAA	ACG	AAA	AAC	CAC	ATT	TAT	CTA	GGT	TTT	AGG	CTA	GGC	ACA	206
Phe	Phe	Lys	Thr	Lys	Asn	His	Ile	Tyr	Leu	Gly	Phe	Arg	Leu	Gly	Thr	
			35					40					45			
GGA	GCC	AAT	GTG	CAC	ACG	AGC	ATG	TGG	CAA	CAA	GCC	TAT	AAA	GAC	AAC	254
Gly	Ala	Asn	Val	His	Thr	Ser	Met	Trp	Gln	Gln	Ala	Tyr	Lys	Asp	Asn	
		50					55					60				
CCC	ACC	TGC	CCT	GGT	AGC	GTG	TGT	TAT	GGC	GAG	AAA	TTA	GAA	GCC	CAT	302
Pro	Thr	Cys	Pro	Gly	Ser	Val	Cys	Tyr	Gly	Glu	Lys	Leu	Glu	Ala	His	
	65					70					75					

- 92 -

TAT CAG GGG GGT AAA AAC CTG TCT TAT ACC GGG CAA ATA GGC GAT GAA	350
Tyr Gln Gly Gly Lys Asn Leu Ser Tyr Thr Gly Gln Ile Gly Asp Glu	
80 85 90 95	
ATA GCT TTT GAT AAA CAC CAT ATT TTA GGC TTA AGG GTG TGG GGG GAT	398
Ile Ala Phe Asp Lys His His Ile Leu Gly Leu Arg Val Trp Gly Asp	
100 105 110	
GTA GAA TAC GCT AAA GCG CAA TTA GGT CAA AAA GTG GGG GGT AAT ACC	446
Val Glu Tyr Ala Lys Ala Gln Leu Gly Gln Lys Val Gly Gly Asn Thr	
115 120 125	
CTT TTA TCC CAA GCC AAT TAT GAC CCA AAC GCG ATT AAA ACC TAC GAT	494
Leu Leu Ser Gln Ala Asn Tyr Asp Pro Asn Ala Ile Lys Thr Tyr Asp	
130 135 140	
TCT GCT TCA AAC ACT CAA GGC CCT TTA GTT TTG CAA AAA ACC CCA AGC	542
Ser Ala Ser Asn Thr Gln Gly Pro Leu Val Leu Gln Lys Thr Pro Ser	
145 150 155	
CCT CAA AAC TTC CTT TTC AAT AAC GGG CAT TTC ATG GCG TTT GGT TTG	590
Pro Gln Asn Phe Leu Phe Asn Asn Gly His Phe Met Ala Phe Gly Leu	
160 165 170 175	
AAC GTG AAT GTG TTT GTT AAC CTC CCT ATA GAC ACC CTT TTA AAA CTC	638
Asn Val Asn Val Phe Val Asn Leu Pro Ile Asp Thr Leu Leu Lys Leu	
180 185 190	
GCT TTA AAA ACA GAA AAA ATG CTG TTT TTT AAA ATA GGC GTG TTT GGT	686
Ala Leu Lys Thr Glu Lys Met Leu Phe Phe Lys Ile Gly Val Phe Gly	
195 200 205	
GGG GGC GGG GTG GAA TAC GCA ATA TTA TGG AGT CCT AAC TAT CAA AAT	734
Gly Gly Gly Val Glu Tyr Ala Ile Leu Trp Ser Pro Asn Tyr Gln Asn	
210 215 220	
CAA AAC ACG AAA CAA GGC GAT AAA TTT TTT GCA GCG GGT GGG GGG TTT	782
Gln Asn Thr Lys Gln Gly Asp Lys Phe Phe Ala Ala Gly Gly Gly Phe	
225 230 235	
TTT GTG AAT TTT GGG GGT TCT TTG TAT ATA GGC AAA CGC AAC CGC TTC	830
Phe Val Asn Phe Gly Gly Ser Leu Tyr Ile Gly Lys Arg Asn Arg Phe	
240 245 250 255	
AAT GTG GGG TTA AAA ATC CCT TAC TAT AGC TTG AGC GCG CAA AGT TGG	878
Asn Val Gly Leu Lys Ile Pro Tyr Tyr Ser Leu Ser Ala Gln Ser Trp	
260 265 270	
AAA AAC TTT GGC TCT AGC AAT GTG TGG CAG CAA CAA ACG ATC CGA CAA	926
Lys Asn Phe Gly Ser Ser Asn Val Trp Gln Gln Gln Thr Ile Arg Gln	
275 280 285	
AAC TTC AGC GTT TTT AGG AAT AAA GAA GTT TTT GTC AGC TAC GCG TTC	974
Asn Phe Ser Val Phe Arg Asn Lys Glu Val Phe Val Ser Tyr Ala Phe	
290 295 300	

-93-

TTG TTT TAGTTTGGAT TCGTTCTCAT TAAACACTGA TGATAAAATT CAAAAGATGG TT 1032
 Leu Phe
 305

TTATCGTTAC AAAATTCAAC ATTTTC 1057

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Gly	Arg	Ile	Glu	Ser	Lys	Lys	Arg	Leu	Lys	Ala	Leu	Val	Phe	Leu	1	5	10	15
Ala	Ser	Leu	Gly	Val	Leu	Trp	Gly	Asn	Ser	Ala	Glu	Lys	Thr	Pro	Phe	20	25	30	
Phe	Lys	Thr	Lys	Asn	His	Ile	Tyr	Leu	Gly	Phe	Arg	Leu	Gly	Thr	Gly	35	40	45	
Ala	Asn	Val	His	Thr	Ser	Met	Trp	Gln	Gln	Ala	Tyr	Lys	Asp	Asn	Pro	50	55	60	
Thr	Cys	Pro	Gly	Ser	Val	Cys	Tyr	Gly	Glu	Lys	Leu	Glu	Ala	His	Tyr	65	70	75	80
Gln	Gly	Gly	Lys	Asn	Leu	Ser	Tyr	Thr	Gly	Gln	Ile	Gly	Asp	Glu	Ile	85	90	95	
Ala	Phe	Asp	Lys	His	His	Ile	Leu	Gly	Leu	Arg	Val	Trp	Gly	Asp	Val	100	105	110	
Glu	Tyr	Ala	Lys	Ala	Gln	Leu	Gly	Gln	Lys	Val	Gly	Gly	Asn	Thr	Leu	115	120	125	
Leu	Ser	Gln	Ala	Asn	Tyr	Asp	Pro	Asn	Ala	Ile	Lys	Thr	Tyr	Asp	Ser	130	135	140	
Ala	Ser	Asn	Thr	Gln	Gly	Pro	Leu	Val	Leu	Gln	Lys	Thr	Pro	Ser	Pro	145	150	155	160
Gln	Asn	Phe	Leu	Phe	Asn	Asn	Gly	His	Phe	Met	Ala	Phe	Gly	Leu	Asn	165	170	175	
Val	Asn	Val	Phe	Val	Asn	Leu	Pro	Ile	Asp	Thr	Leu	Leu	Lys	Leu	Ala	180	185	190	
Leu	Lys	Thr	Glu	Lys	Met	Leu	Phe	Phe	Lys	Ile	Gly	Val	Phe	Gly	Gly	195	200	205	
Gly	Gly	Val	Glu	Tyr	Ala	Ile	Leu	Trp	Ser	Pro	Asn	Tyr	Gln	Asn	Gln	210	215	220	
Asn	Thr	Lys	Gln	Gly	Asp	Lys	Phe	Phe	Ala	Ala	Gly	Gly	Gly	Phe	Phe	225	230	235	240
Val	Asn	Phe	Gly	Gly	Ser	Leu	Tyr	Ile	Gly	Lys	Arg	Asn	Arg	Phe	Asn	245	250	255	
Val	Gly	Leu	Lys	Ile	Pro	Tyr	Tyr	Ser	Leu	Ser	Ala	Gln	Ser	Trp	Lys	260	265	270	
Asn	Phe	Gly	Ser	Ser	Asn	Val	Trp	Gln	Gln	Gln	Thr	Ile	Arg	Gln	Asn	275	280	285	

- 94 -

Phe Ser Val Phe Arg Asn Lys Glu Val Phe Val Ser Tyr Ala Phe Leu
 290 295 300
 Phe
 305

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 624 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 77...535
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TATTAGTTGG TTTAATACGC TATAATCTGT GTGCCAACAT TGTGTGGCTC AAATCATTTT	60
TAAAAGGGGT TTTATA ATG GAA AAC AAC GAA AAT CAT GAG AAA TTG AAT GGC	112
Met Glu Asn Asn Glu Asn His Glu Lys Leu Asn Gly	
1 5 10	
GTT TTG CGC AAG TTT TTA GGC GAT GCG TTC ACG CTT GAT GGG AAA GAA	160
Val Leu Arg Lys Phe Leu Gly Asp Ala Phe Thr Leu Asp Gly Lys Glu	
15 20 25	
GGA GGA TTG AAT ATG GAA AAA TTG CGC GAA GCC ATT AAA AAA GAA AAA	208
Gly Gly Leu Asn Met Glu Lys Leu Arg Glu Ala Ile Lys Lys Glu Lys	
30 35 40	
CCA ATC ATG AAT ATT TTG CTC ATG GGA GCT ACT GGG GTG GGT AAA AGC	256
Pro Ile Met Asn Ile Leu Leu Met Gly Ala Thr Gly Val Gly Lys Ser	
45 50 55 60	
TCG CTC ATT AAC GCT CTA TTC GGT AAG GAA GTA GCT AAA GCA GGT GTA	304
Ser Leu Ile Asn Ala Leu Phe Gly Lys Glu Val Ala Lys Ala Gly Val	
65 70 75	
GGA AAA CCC ATC ACT CAG CAT CTT GAA AAA TAT GTT GAT GAA GAA AAA	352
Gly Lys Pro Ile Thr Gln His Leu Glu Lys Tyr Val Asp Glu Glu Lys	
80 85 90	
GGC TTG ATT TTA TGG GAC ACT AAA GGC ATT GAA GAT AAA GAT TAT GAA	400
Gly Leu Ile Leu Trp Asp Thr Lys Gly Ile Glu Asp Lys Asp Tyr Glu	
95 100 105	
AAT ACC TTG GAA AGC ATT AAA AAA GAA ATG GAA GAT TCT TTT AAA ACG	448
Asn Thr Leu Glu Ser Ile Lys Lys Glu Met Glu Asp Ser Phe Lys Thr	
110 115 120	

- 95 -

CTT GAT GAA AAA GAG GCT ATT GAT GTG GCG TAT CTG TGC GTT AAA GAG 496
 Leu Asp Glu Lys Glu Ala Ile Asp Val Ala Tyr Leu Cys Val Lys Glu
 125 130 135 140

ACT TCT GGT AGG GTT CAA GAG AGA GAG AGA GAG AGT TAT TAAGCTTTAC TA 547
 Thr Ser Gly Arg Val Gln Glu Arg Glu Arg Glu Ser Tyr
 145 150

AAAAATGGAA TATCCCAACG ATTTTCGTTT TCACCAACAC ACAAGAAAAA GCCGGCGATG 607
 CCTTTGTTAA AAAA ACT 624

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Glu Asn Asn Glu Asn His Glu Lys Leu Asn Gly Val Leu Arg Lys
 1 5 10 15
 Phe Leu Gly Asp Ala Phe Thr Leu Asp Gly Lys Glu Gly Gly Leu Asn
 20 25 30
 Met Glu Lys Leu Arg Glu Ala Ile Lys Lys Glu Lys Pro Ile Met Asn
 35 40 45
 Ile Leu Leu Met Gly Ala Thr Gly Val Gly Lys Ser Ser Leu Ile Asn
 50 55 60
 Ala Leu Phe Gly Lys Glu Val Ala Lys Ala Gly Val Gly Lys Pro Ile
 65 70 75 80
 Thr Gln His Leu Glu Lys Tyr Val Asp Glu Lys Gly Leu Ile Leu
 85 90 95
 Trp Asp Thr Lys Gly Ile Glu Asp Lys Asp Tyr Glu Asn Thr Leu Glu
 100 105 110
 Ser Ile Lys Lys Glu Met Glu Asp Ser Phe Lys Thr Leu Asp Glu Lys
 115 120 125
 Glu Ala Ile Asp Val Ala Tyr Leu Cys Val Lys Glu Thr Ser Gly Arg
 130 135 140
 Val Gln Glu Arg Glu Arg Glu Ser Tyr
 145 150

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- 96 -

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 155...1033
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GATGTTGTTA AGTCGTTGTT TATTATGTTA CACTAAAAGC TTAAATAAAA GGGCATAAGG	60
GATAAAGGGA GTGTTAGTAG ATAGTTTAA TAGGGTTATT GACTATATTA GGGTTTCTGT	120
AACCAAACAG TGCAATTTCA GGTGTCAGTA TTGC ATG CCT GCT ACG CCA TTA AAT	175
Met Pro Ala Thr Pro Leu Asn	
1 5	
TTT TTT GAT AAT GAA GAA TTA TTG CCT TTG GAT AAT GTT TTA GAA TTT	223
Phe Phe Asp Asn Glu Glu Leu Leu Pro Leu Asp Asn Val Leu Glu Phe	
10 15 20	
CTC AAA ATC GCC ATT GAT GAG GGC GTT AAA AAA ATT AGA ATC ACG GGT	271
Leu Lys Ile Ala Ile Asp Glu Gly Val Lys Lys Ile Arg Ile Thr Gly	
25 30 35	
GGG GAG CCG CTA TTA CGC AAA GGC TTA GAT GAA TTT ATC GCT AAA TTG	319
Gly Glu Pro Leu Leu Arg Lys Gly Leu Asp Glu Phe Ile Ala Lys Leu	
40 45 50 55	
CAC GCT TAC AAT AAA GAA GTG GAG TTA GTT TTA AGC ACT AAT GGT TTT	367
His Ala Tyr Asn Lys Glu Val Glu Leu Val Leu Ser Thr Asn Gly Phe	
60 65 70	
TTA CTC AAA AAA ATG GCT AAG GAT TTA AAA AAT GCC GGG TTA GCG CAA	415
Leu Leu Lys Lys Met Ala Lys Asp Leu Lys Asn Ala Gly Leu Ala Gln	
75 80 85	
GTG AAT GTT TCA TTG GAT TCT TTA AAA AGC GAT AGG GTT TTA AAA ATC	463
Val Asn Val Ser Leu Asp Ser Leu Lys Ser Asp Arg Val Leu Lys Ile	
90 95 100	
TCT CAA AAA GAC GCT CTT AAA AAC ACG CTA GAA GGG ATT GAA GAG TCT	511
Ser Gln Lys Asp Ala Leu Lys Asn Thr Leu Glu Gly Ile Glu Glu Ser	
105 110 115	
TTG AAA GTG GGT TTA AAA CTC AAA TTA AAC ACG GTT GTG ATA AAA AGC	559
Leu Lys Val Gly Leu Lys Leu Lys Leu Asn Thr Val Val Ile Lys Ser	
120 125 130 135	
GTT AAT GAT GAT GAA ATC TTA GAG CTT TTA GAA TAC GCA AAA AAT AGG	607
Val Asn Asp Asp Glu Ile Leu Glu Leu Leu Glu Tyr Ala Lys Asn Arg	
140 145 150	
CAT ATA CAA ATC CGC TAC ATT GAA TTT ATG GAA AAC ACG CAT GCT AAA	655
His Ile Gln Ile Arg Tyr Ile Glu Phe Met Glu Asn Thr His Ala Lys	
155 160 165	
AGT TTG GTT AAA GGC TTG AAA GAG CGA GAA ATT TTA GAT TTG ATC GCT	703

-97-

Ser	Leu	Val	Lys	Gly	Leu	Lys	Glu	Arg	Glu	Ile	Leu	Asp	Leu	Ile	Ala		
		170					175					180					
CAA	AAA	TAT	CAA	ATC	ATT	GAG	GCA	GAA	AAA	CCC	AAA	CAA	GGG	TCT	TCT	751	
Gln	Lys	Tyr	Gln	Ile	Ile	Glu	Ala	Glu	Lys	Pro	Lys	Gln	Gly	Ser	Ser		
	185					190					195						
AAA	ATC	TAC	ACG	CTA	GAA	AAT	GGC	TAT	CAA	TTT	GGC	ATT	ATC	GCT	CCG	799	
Lys	Ile	Tyr	Thr	Leu	Glu	Asn	Gly	Tyr	Gln	Phe	Gly	Ile	Ile	Ala	Pro		
200					205				210					215			
CAT	AGC	GAT	GAT	TTT	TGC	CAA	TCT	TGC	AAT	CGT	ATC	CGT	TTG	GCT	TCT	847	
His	Ser	Asp	Asp	Phe	Cys	Gln	Ser	Cys	Asn	Arg	Ile	Arg	Leu	Ala	Ser		
				220					225					230			
GAT	GGT	AAG	ATT	TGC	CCA	TGT	TTA	TAC	TAT	CAA	GAC	GCC	ATA	GAC	GCT	895	
Asp	Gly	Lys	Ile	Cys	Pro	Cys	Leu	Tyr	Tyr	Gln	Asp	Ala	Ile	Asp	Ala		
			235					240					245				
AAA	GAG	GCG	ATC	ATC	AAT	AAG	GAT	ACA	AAA	AAT	ATA	AAA	AGG	CTT	TTA	943	
Lys	Glu	Ala	Ile	Ile	Asn	Lys	Asp	Thr	Lys	Asn	Ile	Lys	Arg	Leu	Leu		
		250				255						260					
AAG	CAA	TCT	GTC	ATC	AAT	AAA	CCA	GAA	AAA	AAC	ATG	TGG	AAT	GAT	AAA	991	
Lys	Gln	Ser	Val	Ile	Asn	Lys	Pro	Glu	Lys	Asn	Met	Trp	Asn	Asp	Lys		
	265					270					275						
AAC	AGC	GAA	ACT	CCC	ACA	AGG	GCG	TTT	TAC	TAC	ACA	GGG	GGG	TAGGGGAGT	1042		
Asn	Ser	Glu	Thr	Pro	Thr	Arg	Ala	Phe	Tyr	Tyr	Thr	Gly	Gly				
280						285				290							
AAAATATTTA	TTATTTTAAA	CCTTTTTTATT	AAAAATAAGG	C												1083	

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 293 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Pro	Ala	Thr	Pro	Leu	Asn	Phe	Phe	Asp	Asn	Glu	Glu	Leu	Leu	Pro
1				5					10					15	
Leu	Asp	Asn	Val	Leu	Glu	Phe	Leu	Lys	Ile	Ala	Ile	Asp	Glu	Gly	Val
		20						25					30		
Lys	Lys	Ile	Arg	Ile	Thr	Gly	Gly	Glu	Pro	Leu	Leu	Arg	Lys	Gly	Leu
		35				40						45			
Asp	Glu	Phe	Ile	Ala	Lys	Leu	His	Ala	Tyr	Asn	Lys	Glu	Val	Glu	Leu
	50					55				60					
Val	Leu	Ser	Thr	Asn	Gly	Phe	Leu	Leu	Lys	Lys	Met	Ala	Lys	Asp	Leu

-98-

65		70		75		80									
Lys	Asn	Ala	Gly	Leu	Ala	Gln	Val	Asn	Val	Ser	Leu	Asp	Ser	Leu	Lys
		85						90						95	
Ser	Asp	Arg	Val	Leu	Lys	Ile	Ser	Gln	Lys	Asp	Ala	Leu	Lys	Asn	Thr
		100						105						110	
Leu	Glu	Gly	Ile	Glu	Glu	Ser	Leu	Lys	Val	Gly	Leu	Lys	Leu	Lys	Leu
		115						120						125	
Asn	Thr	Val	Val	Ile	Lys	Ser	Val	Asn	Asp	Asp	Glu	Ile	Leu	Glu	Leu
		130						135						140	
Leu	Glu	Tyr	Ala	Lys	Asn	Arg	His	Ile	Gln	Ile	Arg	Tyr	Ile	Glu	Phe
		145						150						155	160
Met	Glu	Asn	Thr	His	Ala	Lys	Ser	Leu	Val	Lys	Gly	Leu	Lys	Glu	Arg
				165						170					175
Glu	Ile	Leu	Asp	Leu	Ile	Ala	Gln	Lys	Tyr	Gln	Ile	Ile	Glu	Ala	Glu
		180						185						190	
Lys	Pro	Lys	Gln	Gly	Ser	Ser	Lys	Ile	Tyr	Thr	Leu	Glu	Asn	Gly	Tyr
		195						200						205	
Gln	Phe	Gly	Ile	Ile	Ala	Pro	His	Ser	Asp	Asp	Phe	Cys	Gln	Ser	Cys
		210						215						220	
Asn	Arg	Ile	Arg	Leu	Ala	Ser	Asp	Gly	Lys	Ile	Cys	Pro	Cys	Leu	Tyr
		225						230						235	240
Tyr	Gln	Asp	Ala	Ile	Asp	Ala	Lys	Glu	Ala	Ile	Ile	Asn	Lys	Asp	Thr
				245						250					255
Lys	Asn	Ile	Lys	Arg	Leu	Leu	Lys	Gln	Ser	Val	Ile	Asn	Lys	Pro	Glu
				260						265				270	
Lys	Asn	Met	Trp	Asn	Asp	Lys	Asn	Ser	Glu	Thr	Pro	Thr	Arg	Ala	Phe
		275						280						285	
Tyr	Tyr	Thr	Gly	Gly											
		290													

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1181 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 121...1137
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACTTCTCAAT	CAGCGAGCTA	TCATGCAAGG	CCTTATGTGG	TGGATACCGC	TTTTTTACGA	60
TACGATTACA	AAGATGTTTT	TGGGTTTAAG	GCGGGGCGCT	ATGAAGCGAA	TATTGATTTC	120
ATG AGC GGA	TCG AAT CAA	GGG TGG GAA	GTG TAT TAT	CAG CCC	TAT AAG	168
Met Ser Gly	Ser Asn Gln	Gly Trp Glu	Val Tyr Tyr	Gln Pro Tyr	Lys	
1	5	10	15			
ACT GAA	ACG CAA	AGG TTA	AGG TTT	TGG TGG	TGG AGT	216
				TCT TTT	GGG AGA	

-99-

Thr	Glu	Thr	Gln	Arg	Leu	Arg	Phe	Trp	Trp	Trp	Ser	Ser	Phe	Gly	Arg	
			20					25					30			
GGT	TTA	GCG	TTC	AAC	TCT	TGG	ATT	TAT	GAG	TTT	TTT	GCG	ACG	GTG	CCT	264
Gly	Leu	Ala	Phe	Asn	Ser	Trp	Ile	Tyr	Glu	Phe	Phe	Ala	Thr	Val	Pro	
		35					40					45				
TAT	TTG	AAA	AAG	GGA	GGC	AAT	CCT	AAT	AAC	AGC	AAC	GAT	TTC	ATC	AAT	312
Tyr	Leu	Lys	Lys	Gly	Gly	Asn	Pro	Asn	Asn	Ser	Asn	Asp	Phe	Ile	Asn	
	50					55					60					
TAT	GGC	TGG	CAT	GGA	ATC	ACC	ACA	ACC	TAT	TCT	TAT	AAA	GGT	TTA	GAC	360
Tyr	Gly	Trp	His	Gly	Ile	Thr	Thr	Thr	Tyr	Ser	Tyr	Lys	Gly	Leu	Asp	
65					70					75					80	
GCT	CAA	TTT	TTT	TAT	TAT	TTT	GCG	CCT	AAG	ACT	TAT	AAC	GCT	CCT	GGC	408
Ala	Gln	Phe	Phe	Tyr	Tyr	Phe	Ala	Pro	Lys	Thr	Tyr	Asn	Ala	Pro	Gly	
				85					90					95		
TTT	AAG	CTG	GTC	TAT	GAC	ACG	AAT	AGG	AAT	TTT	CAA	AAT	GTA	GGC	TTT	456
Phe	Lys	Leu	Val	Tyr	Asp	Thr	Asn	Arg	Asn	Phe	Gln	Asn	Val	Gly	Phe	
			100					105					110			
CGC	TCT	CAA	AGC	ATG	ATC	ATG	ACA	ACC	TTT	CCT	TTA	TAC	TAT	AGA	GGG	504
Arg	Ser	Gln	Ser	Met	Ile	Met	Thr	Thr	Phe	Pro	Leu	Tyr	Tyr	Arg	Gly	
		115					120					125				
TGG	TAT	AAC	CCA	GAG	ACA	AAC	ACT	TAT	AGT	TTA	GAA	GAC	AGC	ACG	CCT	552
Trp	Tyr	Asn	Pro	Glu	Thr	Asn	Thr	Tyr	Ser	Leu	Glu	Asp	Ser	Thr	Pro	
	130					135					140					
CAT	GGC	TCG	TTG	TTG	GGG	AGG	AAT	GGC	GTT	ACT	TTA	AAT	ATC	CGC	CAG	600
His	Gly	Ser	Leu	Leu	Gly	Arg	Asn	Gly	Val	Thr	Leu	Asn	Ile	Arg	Gln	
145					150				155						160	
GTT	TTT	TGG	TGG	GAT	AAT	TTC	AAC	TGG	TCC	ATT	GGC	TTT	TAT	AAC	ACC	648
Val	Phe	Trp	Trp	Asp	Asn	Phe	Asn	Trp	Ser	Ile	Gly	Phe	Tyr	Asn	Thr	
				165					170					175		
TTT	GGC	AAT	TCG	GAC	GCT	TTT	TTA	GGC	TCT	CAC	ACG	ATG	CCA	AGG	GGT	696
Phe	Gly	Asn	Ser	Asp	Ala	Phe	Leu	Gly	Ser	His	Thr	Met	Pro	Arg	Gly	
			180					185					190			
AAT	AAC	ACT	TCC	TAT	ATC	GGT	AGT	GAA	ATC	TCC	ATA	ACG	ACT	AGG	CAT	744
Asn	Asn	Thr	Ser	Tyr	Ile	Gly	Ser	Glu	Ile	Ser	Ile	Thr	Thr	Arg	His	
		195					200					205				
GCC	GGA	ATG	ATT	GGC	TAT	GAT	TTT	TGG	GAT	AAT	ACG	GCT	TAT	GAT	GGG	792
Ala	Gly	Met	Ile	Gly	Tyr	Asp	Phe	Trp	Asp	Asn	Thr	Ala	Tyr	Asp	Gly	
	210					215					220					
CTA	GCT	GAT	GCG	ATC	ACT	AAC	GCT	AAC	ACT	TTC	ACT	TTT	TAC	ACT	TCT	840
Leu	Ala	Asp	Ala	Ile	Thr	Asn	Ala	Asn	Thr	Phe	Thr	Phe	Tyr	Thr	Ser	
225					230					235					240	

-100-

GTT GGA GGG ATC CAT AAG CGT TTT GCA TGG CAT GTT TTT GGG CGC GTC	888
Val Gly Gly Ile His Lys Arg Phe Ala Trp His Val Phe Gly Arg Val	
245 250 255	
TCT CAT GCG AAT AAA AAC GCG TTA GGG CAA GTG GGG AGG GCT AAT GAA	936
Ser His Ala Asn Lys Asn Ala Leu Gly Gln Val Gly Arg Ala Asn Glu	
260 265 270	
TAT TCC TTG CAA TTC AAC GCG AGC TAT GCG TTC ACT GAA TCA ATC CTT	984
Tyr Ser Leu Gln Phe Asn Ala Ser Tyr Ala Phe Thr Glu Ser Ile Leu	
275 280 285	
CTT AAC TTT AGG ATC ACT TAT TAT GGG GCT AGG ATC AAT AAA GGG TAT	1032
Leu Asn Phe Arg Ile Thr Tyr Tyr Gly Ala Arg Ile Asn Lys Gly Tyr	
290 295 300	
CAA GCG GGG TAT TTT GGA GCG CCC AAA TTC AAT AAC CCT GAT GGC GAT	1080
Gln Ala Gly Tyr Phe Gly Ala Pro Lys Phe Asn Asn Pro Asp Gly Asp	
305 310 315 320	
TTT AGC GCT AAT TAC CAA GAC AGA AGT TAC ATG ATG ACC AAC CTC ACG	1128
Phe Ser Ala Asn Tyr Gln Asp Arg Ser Tyr Met Met Thr Asn Leu Thr	
325 330 335	
CTG AAG TTT TGATTTCCAA TCACAGCGAG TTAAAAACAC TCCAAGGCAT TTTT	1181
Leu Lys Phe	

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ser Gly Ser Asn Gln Gly Trp Glu Val Tyr Tyr Gln Pro Tyr Lys	
1 5 10 15	
Thr Glu Thr Gln Arg Leu Arg Phe Trp Trp Trp Ser Ser Phe Gly Arg	
20 25 30	
Gly Leu Ala Phe Asn Ser Trp Ile Tyr Glu Phe Phe Ala Thr Val Pro	
35 40 45	
Tyr Leu Lys Lys Gly Gly Asn Pro Asn Asn Ser Asn Asp Phe Ile Asn	
50 55 60	
Tyr Gly Trp His Gly Ile Thr Thr Thr Tyr Ser Tyr Lys Gly Leu Asp	
65 70 75 80	
Ala Gln Phe Phe Tyr Phe Ala Pro Lys Thr Tyr Asn Ala Pro Gly	
85 90 95	
Phe Lys Leu Val Tyr Asp Thr Asn Arg Asn Phe Gln Asn Val Gly Phe	
100 105 110	

-101-

```

Arg Ser Gln Ser Met Ile Met Thr Thr Phe Pro Leu Tyr Tyr Arg Gly
      115                      120                      125
Trp Tyr Asn Pro Glu Thr Asn Thr Tyr Ser Leu Glu Asp Ser Thr Pro
      130                      135                      140
His Gly Ser Leu Leu Gly Arg Asn Gly Val Thr Leu Asn Ile Arg Gln
      145                      150                      155                      160
Val Phe Trp Trp Asp Asn Phe Asn Trp Ser Ile Gly Phe Tyr Asn Thr
      165                      170                      175
Phe Gly Asn Ser Asp Ala Phe Leu Gly Ser His Thr Met Pro Arg Gly
      180                      185                      190
Asn Asn Thr Ser Tyr Ile Gly Ser Glu Ile Ser Ile Thr Thr Arg His
      195                      200                      205
Ala Gly Met Ile Gly Tyr Asp Phe Trp Asp Asn Thr Ala Tyr Asp Gly
      210                      215                      220
Leu Ala Asp Ala Ile Thr Asn Ala Asn Thr Phe Thr Phe Tyr Thr Ser
      225                      230                      235                      240
Val Gly Gly Ile His Lys Arg Phe Ala Trp His Val Phe Gly Arg Val
      245                      250                      255
Ser His Ala Asn Lys Asn Ala Leu Gly Gln Val Gly Arg Ala Asn Glu
      260                      265                      270
Tyr Ser Leu Gln Phe Asn Ala Ser Tyr Ala Phe Thr Glu Ser Ile Leu
      275                      280                      285
Leu Asn Phe Arg Ile Thr Tyr Tyr Gly Ala Arg Ile Asn Lys Gly Tyr
      290                      295                      300
Gln Ala Gly Tyr Phe Gly Ala Pro Lys Phe Asn Asn Pro Asp Gly Asp
      305                      310                      315                      320
Phe Ser Ala Asn Tyr Gln Asp Arg Ser Tyr Met Met Thr Asn Leu Thr
      325                      330                      335
Leu Lys Phe

```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 959 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 133...879
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

TAAGGAAATG AGTTTTTATA TCATAAAATA AAGTAACCGA GAAAAATCTT TCTCTAAAAA      60
TAATACTTTT TTAGTTATAA TAACAATTTT GTTTTTTCAA AAACAATAAT TACTATATTT      120
AGGATTTTAA GA ATG AAT GAC AAG CGT TTT AGA AAA TAT TGT AGT TTT TCT      171
      Met Asn Asp Lys Arg Phe Arg Lys Tyr Cys Ser Phe Ser
              1              5              10

```

-102-

ATT TTT TTG TCC TTA TTA GGA ACG TTT GAA TTA GAG GCT AAA GAA GAA	219
Ile Phe Leu Ser Leu Leu Gly Thr Phe Glu Leu Glu Ala Lys Glu Glu	
15 20 25	
GAA GAA AAA GAA GAA AGA AAG ACA GAA AGG AAA AAA GAA AAG AAC GCC	267
Glu Glu Lys Glu Glu Arg Lys Thr Glu Arg Lys Lys Glu Lys Asn Ala	
30 35 40 45	
CAA CAC ACT CTA GGC AAG GTT ACC ACT CAA GCG GCT AAA ATC TTT AAC	315
Gln His Thr Leu Gly Lys Val Thr Thr Gln Ala Ala Lys Ile Phe Asn	
50 55 60	
TAC AAC AAC CAG ACA ACC ATT TCA AGT AAG GAA TTA GAA AGA AGG CAA	363
Tyr Asn Asn Gln Thr Thr Ile Ser Ser Lys Glu Leu Glu Arg Arg Gln	
65 70 75	
GCC AAC CAA ATC AGC GAC ATG TTT AGA AGA AAC CCT AAT ATC AAT GTG	411
Ala Asn Gln Ile Ser Asp Met Phe Arg Arg Asn Pro Asn Ile Asn Val	
80 85 90	
GGC GGT GGT GCG GTG ATA GCG CAA AAA ATT TAT GTG CGC GGT ATT GAA	459
Gly Gly Gly Ala Val Ile Ala Gln Lys Ile Tyr Val Arg Gly Ile Glu	
95 100 105	
GAC AGA TTG GCT CGG GTT ACG GTG GAT GGG GCG GCG CAA ATG GGT GCA	507
Asp Arg Leu Ala Arg Val Thr Val Asp Gly Ala Ala Gln Met Gly Ala	
110 115 120 125	
AGC TAT GGG CAT CAA GGC AAT ACG ATC ATT GAC CCT GGA ATG CTT AAA	555
Ser Tyr Gly His Gln Gly Asn Thr Ile Ile Asp Pro Gly Met Leu Lys	
130 135 140	
AGC GTG GTG GTT ACT AAA GGG GCG GCT CAA GCG AGC GCG GGG CCT ATG	603
Ser Val Val Val Thr Lys Gly Ala Ala Gln Ala Ser Ala Gly Pro Met	
145 150 155	
GCT TTG ATT GGC GCG ATT AAA ATG GAG ACT AAA AGT GCT AGC GAT TTT	651
Ala Leu Ile Gly Ala Ile Lys Met Glu Thr Lys Ser Ala Ser Asp Phe	
160 165 170	
ATC CCT AAA GGT AAA GAC TAC GCC ATA AGT GGG GCT GCC ACT TTT TTA	699
Ile Pro Lys Gly Lys Asp Tyr Ala Ile Ser Gly Ala Ala Thr Phe Leu	
175 180 185	
ACC AAC TTT GGG GAT CGA GAA ACC GTG ATG GGC GCT TAT CGT CAT AAT	747
Thr Asn Phe Gly Asp Arg Glu Thr Val Met Gly Ala Tyr Arg His Asn	
190 195 200 205	
CAT TTT GAT GCG CTT TTG TAT TAC ACG CAT CAA AAT ATT TTT TAC TAT	795
His Phe Asp Ala Leu Leu Tyr Tyr Thr His Gln Asn Ile Phe Tyr Tyr	
210 215 220	
CGT GAT GGG GAT AAT GCT ACA AAA GAT CTC TTT AGA CCT AAA GCG GAG	843
Arg Asp Gly Asp Asn Ala Thr Lys Asp Leu Phe Arg Pro Lys Ala Glu	
225 230 235	

-103-

AAT AAA GTT ACA GAA GTC CTA GCG AGC AAA ACA ATG TGATGGCTAA GATCAA 895
 Asn Lys Val Thr Glu Val Leu Ala Ser Lys Thr Met
 240 245

TGGTTATTTG AGCGAAAGGG ATATTTTAAC GCTCAGTTAT AACATGACCA GAGACAACGC 955
 TAAC 959

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 249 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met	Asn	Asp	Lys	Arg	Phe	Arg	Lys	Tyr	Cys	Ser	Phe	Ser	Ile	Phe	Leu
1				5					10					15	
Ser	Leu	Leu	Gly	Thr	Phe	Glu	Leu	Glu	Ala	Lys	Glu	Glu	Glu	Glu	Lys
			20					25					30		
Glu	Glu	Arg	Lys	Thr	Glu	Arg	Lys	Lys	Glu	Lys	Asn	Ala	Gln	His	Thr
			35				40					45			
Leu	Gly	Lys	Val	Thr	Thr	Gln	Ala	Ala	Lys	Ile	Phe	Asn	Tyr	Asn	Asn
	50					55					60				
Gln	Thr	Thr	Ile	Ser	Ser	Lys	Glu	Leu	Glu	Arg	Arg	Gln	Ala	Asn	Gln
65				70						75				80	
Ile	Ser	Asp	Met	Phe	Arg	Arg	Asn	Pro	Asn	Ile	Asn	Val	Gly	Gly	Gly
				85					90					95	
Ala	Val	Ile	Ala	Gln	Lys	Ile	Tyr	Val	Arg	Gly	Ile	Glu	Asp	Arg	Leu
			100					105					110		
Ala	Arg	Val	Thr	Val	Asp	Gly	Ala	Ala	Gln	Met	Gly	Ala	Ser	Tyr	Gly
		115				120						125			
His	Gln	Gly	Asn	Thr	Ile	Ile	Asp	Pro	Gly	Met	Leu	Lys	Ser	Val	Val
	130					135					140				
Val	Thr	Lys	Gly	Ala	Ala	Gln	Ala	Ser	Ala	Gly	Pro	Met	Ala	Leu	Ile
145				150						155				160	
Gly	Ala	Ile	Lys	Met	Glu	Thr	Lys	Ser	Ala	Ser	Asp	Phe	Ile	Pro	Lys
			165						170					175	
Gly	Lys	Asp	Tyr	Ala	Ile	Ser	Gly	Ala	Ala	Thr	Phe	Leu	Thr	Asn	Phe
			180					185					190		
Gly	Asp	Arg	Glu	Thr	Val	Met	Gly	Ala	Tyr	Arg	His	Asn	His	Phe	Asp
	195					200						205			
Ala	Leu	Leu	Tyr	Tyr	Thr	His	Gln	Asn	Ile	Phe	Tyr	Tyr	Arg	Asp	Gly
	210					215						220			
Asp	Asn	Ala	Thr	Lys	Asp	Leu	Phe	Arg	Pro	Lys	Ala	Glu	Asn	Lys	Val
225				230						235				240	
Thr	Glu	Val	Leu	Ala	Ser	Lys	Thr	Met							
				245											

(2) INFORMATION FOR SEQ ID NO:21:

-104-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1306 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 40...1266
- (D) OTHER INFORMATION:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 40...219
- (D) OTHER INFORMATION:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 220...1266
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTTGACAGCT TATCATTTGG CAATAAAACA CCAAAATGA ATG AGT TAC ACA AAA	54
Met Ser Tyr Thr Lys	
-60	
AAA TAC TCA ACA CCA CCC AAC CGG CGT AAA ATG CAA AAC ATT ATC GCT	102
Lys Tyr Ser Thr Pro Pro Asn Arg Arg Lys Met Gln Asn Ile Ile Ala	
-55 -50 -45 -40	
ATT AAA AGA TCC TCT AGA GTC GAC CTG CAG GCA TGC AAG CTA GCT TTC	150
Ile Lys Arg Ser Arg Val Asp Leu Gln Ala Cys Lys Leu Ala Phe	
-35 -30 -25	
GCG AGC TCG AGA TCA CCC ATG CAA TTT CAA AAA ACC TTA TTT CCT TTA	198
Ala Ser Ser Arg Ser Pro Met Gln Phe Gln Lys Thr Leu Phe Pro Leu	
-20 -15 -10	
CCC TTA TTA TTT TTA TCT TGT TGT ATC GCT GAA GAA AAT GGG GCG TAT	246
Pro Leu Leu Phe Leu Ser Cys Cys Ile Ala Glu Glu Asn Gly Ala Tyr	
-5 1 5	
GCG AGC GTG GGG TTT GAA TAT TCC ATT AGT CAT GCC GTT GAG CAT AAT	294
Ala Ser Val Gly Phe Glu Tyr Ser Ile Ser His Ala Val Glu His Asn	
10 15 20 25	
AAC CCT TTT TTA AAT CAA GAA CGC ATC CAA ATC ATT TCT AAC GCT CAA	342
Asn Pro Phe Leu Asn Gln Glu Arg Ile Gln Ile Ile Ser Asn Ala Gln	
30 35 40	
AAC AAA ATC TAT AAA CTC AAT CAA GTC AAA AAT GAA ATC ACA AGC ATG	390

-105-

Asn	Lys	Ile	Tyr	Lys	Leu	Asn	Gln	Val	Lys	Asn	Glu	Ile	Thr	Ser	Met	
			45					50					55			
CAA	AAC	ACC	TTT	AAT	TAC	ATC	AAC	AAC	GCT	TTA	AAA	AAC	AAT	GCT	AAA	438
Gln	Asn	Thr	Phe	Asn	Tyr	Ile	Asn	Asn	Ala	Leu	Lys	Asn	Asn	Ala	Lys	
		60					65					70				
TTA	ACC	CCC	ACT	GAA	ATC	CAA	GCT	GAG	AAA	TAC	TAC	CTC	CAA	TCC	ACC	486
Leu	Thr	Pro	Thr	Glu	Ile	Gln	Ala	Glu	Lys	Tyr	Tyr	Leu	Gln	Ser	Thr	
	75					80					85					
CTT	CAA	AAC	ATT	GAA	AAA	ATA	GTC	ACA	CTT	AGC	GGT	GGC	GTT	GCA	TCT	534
Leu	Gln	Asn	Ile	Glu	Lys	Ile	Val	Thr	Leu	Ser	Gly	Gly	Val	Ala	Ser	
90					95					100					105	
AAC	CCC	AAA	CTA	GTC	CAA	GCG	TTG	GAA	AAA	ATG	CAA	GAA	CCC	ATT	ACT	582
Asn	Pro	Lys	Leu	Val	Gln	Ala	Leu	Glu	Lys	Met	Gln	Glu	Pro	Ile	Thr	
				110					115					120		
AAC	CCT	TTA	GAA	TTA	GCA	GAA	AAC	TTA	AGA	AAT	TTA	GAA	TTG	CAA	TTT	630
Asn	Pro	Leu	Glu	Leu	Ala	Glu	Asn	Leu	Arg	Asn	Leu	Glu	Leu	Gln	Phe	
			125					130					135			
GCT	CAA	TCT	CAA	AAC	CGC	ATG	CTT	TCT	TCT	TTA	TCT	TCT	CAA	ACC	GCT	678
Ala	Gln	Ser	Gln	Asn	Arg	Met	Leu	Ser	Ser	Leu	Ser	Ser	Gln	Thr	Ala	
		140					145						150			
CAA	ATT	TCA	AAT	TCT	TTG	AAC	GCG	CTT	GAT	CCC	AGC	TCT	TAT	TCT	AAA	726
Gln	Ile	Ser	Asn	Ser	Leu	Asn	Ala	Leu	Asp	Pro	Ser	Ser	Tyr	Ser	Lys	
	155					160					165					
AAC	ATT	TCA	AGC	ATG	TCT	GGG	GTG	AGT	TTG	AGC	GTA	GGG	TAT	AAG	CAT	774
Asn	Ile	Ser	Ser	Met	Ser	Gly	Val	Ser	Leu	Ser	Val	Gly	Tyr	Lys	His	
170					175					180					185	
TTC	TTT	ACT	AAG	AAA	AAA	AAT	CAA	GGG	TTT	CGC	TAT	TAC	TTG	TTT	TAT	822
Phe	Phe	Thr	Lys	Lys	Lys	Asn	Gln	Gly	Phe	Arg	Tyr	Tyr	Leu	Phe	Tyr	
				190					195					200		
GAC	TAT	GGT	TAC	ACT	AAC	TTT	GGT	TTT	GTG	GGT	AAT	GGC	TTT	GAT	GGT	870
Asp	Tyr	Gly	Tyr	Thr	Asn	Phe	Gly	Phe	Val	Gly	Asn	Gly	Phe	Asp	Gly	
			205					210					215			
TTA	GGC	AAA	ATG	AAT	AAC	CAC	CTC	TAT	GGG	CTT	GGA	ATA	AAC	TAC	CTT	918
Leu	Gly	Lys	Met	Asn	Asn	His	Leu	Tyr	Gly	Leu	Gly	Ile	Asn	Tyr	Leu	
		220					225					230				
TAT	AAT	TTC	ATT	GAT	AAT	GCA	CAA	AAA	CAT	TCG	AGC	GTG	GGT	TTT	TAT	966
Tyr	Asn	Phe	Ile	Asp	Asn	Ala	Gln	Lys	His	Ser	Ser	Val	Gly	Phe	Tyr	
	235					240					245					
GCG	GGT	TTT	GCT	TTG	GCG	GGG	AAT	TCG	TGG	GTA	GGG	AAT	GGT	TTA	GGC	1014
Ala	Gly	Phe	Ala	Leu	Ala	Gly	Asn	Ser	Trp	Val	Gly	Asn	Gly	Leu	Gly	
250					255					260					265	

-106-

ATG TGG GTG AGC CAA ACG GAT TTT ATC AAC AAT TAC TTG ATG GGC TAT	1062
Met Trp Val Ser Gln Thr Asp Phe Ile Asn Asn Tyr Leu Met Gly Tyr	
270 275 280	
CAA GCT AAA ATA CAC ACG AAC TTT TTC CAG ATC CCT TTG AAT TTT GGG	1110
Gln Ala Lys Ile His Thr Asn Phe Phe Gln Ile Pro Leu Asn Phe Gly	
285 290 295	
GTT CGT GTG AAT GTC AAT AGG CAT AAC GGA TTT GAA ATG GGC CTA AAA	1158
Val Arg Val Asn Val Asn Arg His Asn Gly Phe Glu Met Gly Leu Lys	
300 305 310	
ATC CCT TTA GCG GTG AAT TCC TTT TAT GAA ACG CAT GGC AAA GGG TTA	1206
Ile Pro Leu Ala Val Asn Ser Phe Tyr Glu Thr His Gly Lys Gly Leu	
315 320 325	
AAC ACT TCC CTC TTT TTC AAA CGC CTT GTG GTG TTT AAT GTG AGT TAT	1254
Asn Thr Ser Leu Phe Phe Lys Arg Leu Val Val Phe Asn Val Ser Tyr	
330 335 340 345	
GTT TAT AGT TTT TAGGGGGTAA ATGCCTTCAA ACGCTCTTTT GATTGAAGAA	1306
Val Tyr Ser Phe	

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 409 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Ser Tyr Thr Lys Lys Tyr Ser Thr Pro Pro Asn Arg Arg Lys Met	
-60 -55 -50 -45	
Gln Asn Ile Ile Ala Ile Lys Arg Ser Ser Arg Val Asp Leu Gln Ala	
-40 -35 -30	
Cys Lys Leu Ala Phe Ala Ser Ser Arg Ser Pro Met Gln Phe Gln Lys	
-25 -20 -15	
Thr Leu Phe Pro Leu Pro Leu Leu Phe Leu Ser Cys Cys Ile Ala Glu	
-10 -5 1	
Glu Asn Gly Ala Tyr Ala Ser Val Gly Phe Glu Tyr Ser Ile Ser His	
5 10 15 20	
Ala Val Glu His Asn Asn Pro Phe Leu Asn Gln Glu Arg Ile Gln Ile	
25 30 35	
Ile Ser Asn Ala Gln Asn Lys Ile Tyr Lys Leu Asn Gln Val Lys Asn	
40 45 50	
Glu Ile Thr Ser Met Gln Asn Thr Phe Asn Tyr Ile Asn Asn Ala Leu	
55 60 65	
Lys Asn Asn Ala Lys Leu Thr Pro Thr Glu Ile Gln Ala Glu Lys Tyr	

-107-

70		75		80											
Tyr	Leu	Gln	Ser	Thr	Leu	Gln	Asn	Ile	Glu	Lys	Ile	Val	Thr	Leu	Ser
85					90					95					100
Gly	Gly	Val	Ala	Ser	Asn	Pro	Lys	Leu	Val	Gln	Ala	Leu	Glu	Lys	Met
				105						110					115
Gln	Glu	Pro	Ile	Thr	Asn	Pro	Leu	Glu	Leu	Ala	Glu	Asn	Leu	Arg	Asn
			120					125					130		
Leu	Glu	Leu	Gln	Phe	Ala	Gln	Ser	Gln	Asn	Arg	Met	Leu	Ser	Ser	Leu
	135						140					145			
Ser	Ser	Gln	Thr	Ala	Gln	Ile	Ser	Asn	Ser	Leu	Asn	Ala	Leu	Asp	Pro
	150						155				160				
Ser	Ser	Tyr	Ser	Lys	Asn	Ile	Ser	Ser	Met	Ser	Gly	Val	Ser	Leu	Ser
165					170					175					180
Val	Gly	Tyr	Lys	His	Phe	Phe	Thr	Lys	Lys	Lys	Asn	Gln	Gly	Phe	Arg
			185						190					195	
Tyr	Tyr	Leu	Phe	Tyr	Asp	Tyr	Gly	Tyr	Thr	Asn	Phe	Gly	Phe	Val	Gly
		200					205					210			
Asn	Gly	Phe	Asp	Gly	Leu	Gly	Lys	Met	Asn	Asn	His	Leu	Tyr	Gly	Leu
	215						220					225			
Gly	Ile	Asn	Tyr	Leu	Tyr	Asn	Phe	Ile	Asp	Asn	Ala	Gln	Lys	His	Ser
	230					235					240				
Ser	Val	Gly	Phe	Tyr	Ala	Gly	Phe	Ala	Leu	Ala	Gly	Asn	Ser	Trp	Val
245					250					255					260
Gly	Asn	Gly	Leu	Gly	Met	Trp	Val	Ser	Gln	Thr	Asp	Phe	Ile	Asn	Asn
			265						270					275	
Tyr	Leu	Met	Gly	Tyr	Gln	Ala	Lys	Ile	His	Thr	Asn	Phe	Phe	Gln	Ile
		280						285					290		
Pro	Leu	Asn	Phe	Gly	Val	Arg	Val	Asn	Val	Asn	Arg	His	Asn	Gly	Phe
	295						300					305			
Glu	Met	Gly	Leu	Lys	Ile	Pro	Leu	Ala	Val	Asn	Ser	Phe	Tyr	Glu	Thr
	310					315					320				
His	Gly	Lys	Gly	Leu	Asn	Thr	Ser	Leu	Phe	Phe	Lys	Arg	Leu	Val	Val
325				330						335					340
Phe	Asn	Val	Ser	Tyr	Val	Tyr	Ser	Phe							
			345												

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1030 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 342...824
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

-108-

CACTCTAAGC	GTCAAACCTCT	CTTTTTCTTT	AGAGGAAGAA	AGCAAGCGGA	TCCATCTTAA	60
AGCCTTACAA	AATATCTTAA	ATAACGCTAA	AAGCGCGCAT	TTTAAATTTG	TTTTAGAGAG	120
CCAAAACGCC	GCTCAATCTA	TTATAGAAAT	TCAAAGCCTC	TTGAAACAAC	TCTCCTTAAA	180
AAATAATGAA	ATCTTTTTTAA	TGCCTTTAGG	CACAAATAAC	AACGAGCTAG	ACAAAAATCT	240
AAAAACCCTA	GCCCCCCTAG	CCATAAAGCA	TGGTTTCAGG	CTAAGCGATA	GGCTTCATAT	300
CCGCTTGTGG	GATAATCAAA	AAGGGTTTTA	AAAAGTTAAT	C ATG ACC ATC AAA GTT		356
				Met Thr Ile Lys Val		
				1 5		
TTT TCG CCC AAA TAC CCC ACT GAA TTA GAA GAA TTT TAT GCT GAG CGT						404
Phe Ser Pro Lys Tyr Pro Thr Glu Leu Glu Glu Phe Tyr Ala Glu Arg						
	10		15		20	
ATC GCT GAC AAC CCT TTA GGG TTT ATC CAA CGC TTG GAT CTT TTG CCT						452
Ile Ala Asp Asn Pro Leu Gly Phe Ile Gln Arg Leu Asp Leu Leu Pro						
	25		30		35	
AGT ATT AGC GGG TTC GTT CAA AAA TTG CGC GAG CAT GGC GGG GAA TTT						500
Ser Ile Ser Gly Phe Val Gln Lys Leu Arg Glu His Gly Gly Glu Phe						
	40		45		50	
TTT GAA ATG AGA GAG GGT AAC AAG CTC ATT GGG ATT TGT GGG CTT AAT						548
Phe Glu Met Arg Glu Gly Asn Lys Leu Ile Gly Ile Cys Gly Leu Asn						
	55		60		65	
CCT ATC AAT CAA ACA GAA GCC GAG CTG TGC AAA TTC CAC ATA AAT AGT						596
Pro Ile Asn Gln Thr Glu Ala Glu Leu Cys Lys Phe His Ile Asn Ser						
	70		75		80	85
GCT TAT CAA TCC CAA GGG CTA GGT CAA AAA CTC TAT GAG AGC GTG GAG						644
Ala Tyr Gln Ser Gln Gly Leu Gly Gln Lys Leu Tyr Glu Ser Val Glu						
	90		95		100	
AAA TAC GCT TTC ATT AAA GGC TAT ACT AAA ATC TCT CTG CAT GTG AGC						692
Lys Tyr Ala Phe Ile Lys Gly Tyr Thr Lys Ile Ser Leu His Val Ser						
	105		110		115	
AAA AGC CAA ATC AAG GCA TGC AAC CTC TAT CAA AAG CTG GGT TTT GTG						740
Lys Ser Gln Ile Lys Ala Cys Asn Leu Tyr Gln Lys Leu Gly Phe Val						
	120		125		130	
CAC ATC AAA GAA GAG GAT TGC GTG GTG GAG TTG GGC GAA GAG ACT TTG						788
His Ile Lys Glu Glu Asp Cys Val Val Glu Leu Gly Glu Glu Thr Leu						
	135		140		145	
ATT TTC CCC ACT CTT TTT ATG GAA AAG ATT TTG TCT TGATTGGTGC ATCCAT						840
Ile Phe Pro Thr Leu Phe Met Glu Lys Ile Leu Ser						
	150		155		160	
TTGACACACG	CCCAAGCGAC	ATTCAAACCTA	TCAAACCTTTC	ATTAACACAA	CCCAATTAAC	900
GCTAAATAAA	CCCTAAAACA	AACACTCGTT	GTTAAAAATT	TGTTTTTCAA	GCGCTTCGCA	960
AAGTTTTAGA	AGCCCTATTT	AGGGGTTAAC	GCTAAAAATAG	GCTATCAAAA	CTACTTTAAT	1020
GATTTTATAG						1030

(2) INFORMATION FOR SEQ ID NO:24:

-109-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```

Met Thr Ile Lys Val Phe Ser Pro Lys Tyr Pro Thr Glu Leu Glu Glu
 1             5             10             15
Phe Tyr Ala Glu Arg Ile Ala Asp Asn Pro Leu Gly Phe Ile Gln Arg
      20             25             30
Leu Asp Leu Leu Pro Ser Ile Ser Gly Phe Val Gln Lys Leu Arg Glu
      35             40             45
His Gly Gly Glu Phe Phe Glu Met Arg Glu Gly Asn Lys Leu Ile Gly
      50             55             60
Ile Cys Gly Leu Asn Pro Ile Asn Gln Thr Glu Ala Glu Leu Cys Lys
65             70             75             80
Phe His Ile Asn Ser Ala Tyr Gln Ser Gln Gly Leu Gly Gln Lys Leu
      85             90             95
Tyr Glu Ser Val Glu Lys Tyr Ala Phe Ile Lys Gly Tyr Thr Lys Ile
      100            105            110
Ser Leu His Val Ser Lys Ser Gln Ile Lys Ala Cys Asn Leu Tyr Gln
      115            120            125
Lys Leu Gly Phe Val His Ile Lys Glu Glu Asp Cys Val Val Glu Leu
      130            135            140
Gly Glu Glu Thr Leu Ile Phe Pro Thr Leu Phe Met Glu Lys Ile Leu
145            150            155            160
Ser

```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1477 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 374...1267
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

CGTGGAGTTT TTTAGGCATT TCTTTATATT CATTCAATAA CGCTTGCGCG GGCAATTCTT      60
CAACTAAAAT CTCTACTAAC AATTCATCTG AATGCAAAAT CTCAATTCTC CCTAAAAAAC      120
AAAATCACTT TTAAGACTAA ATCATGTTAG AATTATACTT GAATTACAC TCAGTTTAGT      180

```

-110-

TTATTTCTTA ATACAAAAGG TAGGCGTTTT GAAACATTTA ACCCCACTCA CTCACACCAT	240
CTTTAAAGCC TTATGGCTAG GCACAGCCTT AAGTGCATCT TTAAGTTTAG CCGCAACAGA	300
AAGCCCCACT AAAACAGAGC CTAAGCCCCG TAAAGGGGTT AAAAACAAGC CCAAATCGCC	360
CGTTACTAAA GTC ATG ATG ACC AAT TGC GAC AAT ATT AAA GAT TTT AAC	409
Met Met Thr Asn Cys Asp Asn Ile Lys Asp Phe Asn	
1 5 10	
GCT AAG CAA AAA GAA GTC TTA AAA GCC GCT TAT CAA TTC GGC TCT AAA	457
Ala Lys Gln Lys Glu Val Leu Lys Ala Ala Tyr Gln Phe Gly Ser Lys	
15 20 25	
GAA AAT TTA GGC TAT GAA ATG GCA GGC ATT GCA TGG AAA GAG TCA TGC	505
Glu Asn Leu Gly Tyr Glu Met Ala Gly Ile Ala Trp Lys Glu Ser Cys	
30 35 40	
GCA GGG GTT TAT AAA ATC AAT TTT TCG GAT CCG AGC GCG GGC GTG TAT	553
Ala Gly Val Tyr Lys Ile Asn Phe Ser Asp Pro Ser Ala Gly Val Tyr	
45 50 55 60	
CAT TCT TAT ATC CCA AGC GTT CTA AAA AGC TAT GGG CAT AAT GAT AGC	601
His Ser Tyr Ile Pro Ser Val Leu Lys Ser Tyr Gly His Asn Asp Ser	
65 70 75	
CCC TTT TTG CGT AAT GTG ATG GGG GAA TTG CTC ATT AAA GAC GAT GCG	649
Pro Phe Leu Arg Asn Val Met Gly Glu Leu Leu Ile Lys Asp Asp Ala	
80 85 90	
TTT GCT TCT GAA GTG GCT TTA AAA GAG TTG CTC TAT TGG AAA ACA CGC	697
Phe Ala Ser Glu Val Ala Leu Lys Glu Leu Leu Tyr Trp Lys Thr Arg	
95 100 105	
TAC CAT GAC AAT TTA AAA GAC ATG ATT AAA TCT TAC AAC AAG GGC AGT	745
Tyr His Asp Asn Leu Lys Asp Met Ile Lys Ser Tyr Asn Lys Gly Ser	
110 115 120	
CGT TGG GAA AGG AGC GAA AAA TCT AAC GCT GAT GCT GAA AAA TAT TAC	793
Arg Trp Glu Arg Ser Glu Lys Ser Asn Ala Asp Ala Glu Lys Tyr Tyr	
125 130 135 140	
GAA GAG ATA CAA GAC AGA ATC AGG CGT TTG AAA GAA TCT AAA ATC TTT	841
Glu Glu Ile Gln Asp Arg Ile Arg Arg Leu Lys Glu Ser Lys Ile Phe	
145 150 155	
GAT TCG CAG TCT AGT AAT GAC CAA GAA TTG CAA AAA AGC GCT AAT AGC	889
Asp Ser Gln Ser Ser Asn Asp Gln Glu Leu Gln Lys Ser Ala Asn Ser	
160 165 170	
AAC CTG GAT TTA GAC CCT ATC GGC AAC GCC ATG CCC CAA GCC TTA ATT	937
Asn Leu Asp Leu Asp Pro Ile Gly Asn Ala Met Pro Gln Ala Leu Ile	
175 180 185	
GCC AAA GAA ACT AAA ATA GAA GAA ACC CAA GCA GAA AAA TCC CAA GAA	985
Ala Lys Glu Thr Lys Ile Glu Glu Thr Gln Ala Glu Lys Ser Gln Glu	
190 195 200	

-111-

ATG AAA GAG ACA ACT AGC GAG CAA ACA AAA AGT AAG CCA GAA AAA GCA	1033
Met Lys Glu Thr Thr Ser Glu Gln Thr Lys Ser Lys Pro Glu Lys Ala	
205 210 215 220	
AAA GAT AAA CCC ATG TAT TTG GCG CAA ATC AAC AGC ACT GAT TTC ACA	1081
Lys Asp Lys Pro Met Tyr Leu Ala Gln Ile Asn Ser Thr Asp Phe Thr	
225 230 235	
CCC GTT AAA AAA AGC CCC AAA AAA CCG GCT AAA GTG AGC CAA AAA CAC	1129
Pro Val Lys Lys Ser Pro Lys Lys Pro Ala Lys Val Ser Gln Lys His	
240 245 250	
TCC TTT AAG AAT AAC ATT AAA AAT AAT GTA AAA AAC AAC GCC AAA ACC	1177
Ser Phe Lys Asn Asn Ile Lys Asn Asn Val Lys Asn Asn Ala Lys Thr	
255 260 265	
GCT TCC AAA AAA CAA GAA ATG TGC AAA AAT TGC TCT CCA GGG CAA AGG	1225
Ala Ser Lys Lys Gln Glu Met Cys Lys Asn Cys Ser Pro Gly Gln Arg	
270 275 280	
AAT GCG ATT TTA GCT AAC CAC ATC ACT CTC ATG CAA GAG CTT TAAAAAGTC	1276
Asn Ala Ile Leu Ala Asn His Ile Thr Leu Met Gln Glu Leu	
285 290 295	
CTAAAAATGG CGCAAAAAAC TCTTTTGATT ATCACTGATG GCATTGGGTA TCGTAAAGAT	1336
AGCGATCATA ACGCTTTCTT CCATGCCAAA AAACCCACTT ATGATTTGAT GTTTAAACC	1396
TTGCCTTATA GCCTGATTGA TACGCATGGC TTGAGCGTGG GCTTACCTAA GGGGCAAATG	1456
GGAAATTCTG AAGTGGGGCA T	1477

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 298 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Met Thr Asn Cys Asp Asn Ile Lys Asp Phe Asn Ala Lys Gln Lys	
1 5 10 15	
Glu Val Leu Lys Ala Ala Tyr Gln Phe Gly Ser Lys Glu Asn Leu Gly	
20 25 30	
Tyr Glu Met Ala Gly Ile Ala Trp Lys Glu Ser Cys Ala Gly Val Tyr	
35 40 45	
Lys Ile Asn Phe Ser Asp Pro Ser Ala Gly Val Tyr His Ser Tyr Ile	
50 55 60	
Pro Ser Val Leu Lys Ser Tyr Gly His Asn Asp Ser Pro Phe Leu Arg	
65 70 75 80	
Asn Val Met Gly Glu Leu Leu Ile Lys Asp Asp Ala Phe Ala Ser Glu	
85 90 95	
Val Ala Leu Lys Glu Leu Leu Tyr Trp Lys Thr Arg Tyr His Asp Asn	

[illegible]

(A) LENGTH: 1515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 141...1340
(D) OTHER INFORMATION:

TTAGTGTGTA	TTTTTTTATC	GTTAGTGT	TTT	GTGCGTCCTT	TAGAGGCTTT	GAGCGTGT	TTT	60
ATGGGGTTGT	ATTTGATTTA	TGGCATCATT	CGGTGGCTCT	TTTTAATGGT	AAAAATTATT			120
TTTAATAAAAA	ATAAAAGCGC	ATG AAA GAA TCT TTT TAC ATA GAG GGA ATG						170
		Met Lys Glu Ser Phe Tyr Ile Glu Gly Met						
		1		5			10	
ACT TGC ACG GCG TGT TCT AGC GGG ATT GAA CGC TCT TTG GGG CGT AAG								218
Thr Cys Thr Ala Cys Ser Ser Gly Ile Glu Arg Ser Leu Gly Arg Lys								
		15		20			25	
AGT TTT GTG AAA AAA ATA GAA GTG AGC CTT TTA AAT AAG AGC GCT AAC								266
Ser Phe Val Lys Lys Ile Glu Val Ser Leu Leu Asn Lys Ser Ala Asn								

-113-

30				35				40								
ATT	GAA	TTT	GAC	GAA	AAC	CAA	ACC	AAT	TTA	GAC	GAA	ATT	TTT	AAA	CTC	314
Ile	Glu	Phe	Asp	Glu	Asn	Gln	Thr	Asn	Leu	Asp	Glu	Ile	Phe	Lys	Leu	
		45					50						55			
ATT	GAA	AAG	CTA	GGC	TAT	AGC	CCT	AAA	AAA	GCT	CTG	ACA	AAA	GAA	AAA	362
Ile	Glu	Lys	Leu	Gly	Tyr	Ser	Pro	Lys	Lys	Ala	Leu	Thr	Lys	Glu	Lys	
		60				65					70					
AAA	GAA	TTT	TTT	AGC	CCT	AAT	GTT	AAA	TTA	GCG	TTA	GCG	GTT	ATT	TTC	410
Lys	Glu	Phe	Phe	Ser	Pro	Asn	Val	Lys	Leu	Ala	Leu	Ala	Val	Ile	Phe	
		75				80				85					90	
ACG	CTT	TTT	GTG	GTG	TAT	CTT	TCT	ATG	GGG	GCG	ATG	CTT	AGC	CCT	AGC	458
Thr	Leu	Phe	Val	Val	Tyr	Leu	Ser	Met	Gly	Ala	Met	Leu	Ser	Pro	Ser	
				95					100					105		
CTT	TTA	CCT	GAA	AGC	TTG	CTT	GCA	ATT	GAT	AAT	CAT	AGT	AAT	TTT	TTA	506
Leu	Leu	Pro	Glu	Ser	Leu	Leu	Ala	Ile	Asp	Asn	His	Ser	Asn	Phe	Leu	
			110					115					120			
AAC	GCT	TGC	TTA	CAG	CTT	ATA	GGC	GCA	CTC	ATT	GTC	ATG	CAT	TTG	GGG	554
Asn	Ala	Cys	Leu	Gln	Leu	Ile	Gly	Ala	Leu	Ile	Val	Met	His	Leu	Gly	
		125					130					135				
AGG	GAT	TTT	TAC	ATT	CAA	GGG	TTT	AAA	GCC	TTA	TGG	CAC	AGA	CAA	CCC	602
Arg	Asp	Phe	Tyr	Ile	Gln	Gly	Phe	Lys	Ala	Leu	Trp	His	Arg	Gln	Pro	
		140				145					150					
AAC	ATG	AGC	AGC	CTT	ATC	GCC	ATA	GGC	ACA	AGC	GCT	GCC	TTA	ATT	TCA	650
Asn	Met	Ser	Ser	Leu	Ile	Ala	Ile	Gly	Thr	Ser	Ala	Ala	Leu	Ile	Ser	
		155				160				165					170	
AGC	CTG	TGG	CAA	TTG	TAT	TTG	GTC	TAT	ACC	AAT	CAT	TAT	ACC	GAT	CAG	698
Ser	Leu	Trp	Gln	Leu	Tyr	Leu	Val	Tyr	Thr	Asn	His	Tyr	Thr	Asp	Gln	
				175					180					185		
TGG	TCT	TAT	GGG	CAT	TAT	TAT	TTT	GAA	AGC	GTG	TGC	GTG	ATT	TTA	ATG	746
Trp	Ser	Tyr	Gly	His	Tyr	Tyr	Phe	Glu	Ser	Val	Cys	Val	Ile	Leu	Met	
			190					195					200			
TTT	GTG	ATG	GTG	GGC	AAA	CGC	ATT	GAA	AAT	GTT	TCT	AAA	GAC	AAA	GCT	794
Phe	Val	Met	Val	Gly	Lys	Arg	Ile	Glu	Asn	Val	Ser	Lys	Asp	Lys	Ala	
		205					210					215				
TTA	GAC	GCT	ATG	CAA	GCC	TTG	ATG	AAA	AAC	GCC	CCA	AAA	ACC	GCC	CTT	842
Leu	Asp	Ala	Met	Gln	Ala	Leu	Met	Lys	Asn	Ala	Pro	Lys	Thr	Ala	Leu	
		220				225					230					
AAA	ATG	CAA	AAT	AAC	CAA	CAG	ATT	GAA	GTT	TTA	GTG	GAT	AGC	ATT	GTG	890
Lys	Met	Gln	Asn	Asn	Gln	Gln	Ile	Glu	Val	Leu	Val	Asp	Ser	Ile	Val	
		235				240				245					250	
GTG	GGG	GAT	ATT	CTA	AAA	GTC	CTC	CCT	GGA	AGC	GCG	ATT	GCG	GTG	GAT	938

-114-

Val	Gly	Asp	Ile	Leu	Lys	Val	Leu	Pro	Gly	Ser	Ala	Ile	Ala	Val	Asp	
				255					260					265		
GGT	GAA	ATC	ATA	GAG	GGC	GAA	GGG	GAA	TTA	GAT	GAG	AGC	ATG	TTG	AGC	986
Gly	Glu	Ile	Ile	Glu	Gly	Glu	Gly	Glu	Leu	Asp	Glu	Ser	Met	Leu	Ser	
			270					275					280			
GGC	GAA	GCG	TTG	CCG	GTT	TAT	AAA	AAA	GTC	GGC	GAT	AAA	GTC	TTT	TCA	1034
Gly	Glu	Ala	Leu	Pro	Val	Tyr	Lys	Lys	Val	Gly	Asp	Lys	Val	Phe	Ser	
		285					290					295				
GGG	ACA	TTC	AAT	AGC	CAC	ACG	AGT	TTT	TTA	ATG	AAA	GCC	ACG	CAA	AAC	1082
Gly	Thr	Phe	Asn	Ser	His	Thr	Ser	Phe	Leu	Met	Lys	Ala	Thr	Gln	Asn	
	300					305					310					
AAC	AAA	AAC	AGC	ACC	TTG	TCT	CAA	ATT	ATA	GAA	ATG	ATT	TAT	AAC	GCT	1130
Asn	Lys	Asn	Ser	Thr	Leu	Ser	Gln	Ile	Ile	Glu	Met	Ile	Tyr	Asn	Ala	
315					320					325				330		
CAA	AGT	TCA	AAG	GCA	GAG	ATT	TCT	CGC	TTA	GCG	GAT	AAG	GTT	TCA	AGC	1178
Gln	Ser	Ser	Lys	Ala	Glu	Ile	Ser	Arg	Leu	Ala	Asp	Lys	Val	Ser	Ser	
			335						340					345		
GTG	TTT	GTG	CCA	AGC	GTG	ATC	GCT	ATT	TCT	ATT	TTA	GCG	TTT	GTG	GTG	1226
Val	Phe	Val	Pro	Ser	Val	Ile	Ala	Ile	Ser	Ile	Leu	Ala	Phe	Val	Val	
			350					355					360			
TGG	CTC	ATC	ATT	GCA	CCT	AAG	CCC	GAT	TTT	TGG	TGG	AAT	TTT	GGA	ATC	1274
Trp	Leu	Ile	Ile	Ala	Pro	Lys	Pro	Asp	Phe	Trp	Trp	Asn	Phe	Gly	Ile	
	365						370					375				
GCT	TTA	GAA	GTG	TTT	GTA	TCG	GTT	TTA	GTG	ATT	TCT	TGC	CCT	TGC	GCT	1322
Ala	Leu	Glu	Val	Phe	Val	Ser	Val	Leu	Val	Ile	Ser	Cys	Pro	Cys	Ala	
	380					385					390					
TTA	GGA	TTG	CTA	CGC	CTA	TGAGC	ATTTT	AGTAG	CGAAC	CAGAA	AGCGA	GTTCT	TTTA			1378
Leu	Gly	Leu	Leu	Arg	Leu											
395					400											
GGGTTATTTT	TTAAAGACGC	TAAAAGTTTA	GAAAAAGCAA	GGCTAGTCAA	TACGATCGTT											1438
TTTGATAAAA	CCGGCACGCT	CACTAACGGC	AAGCCTGTCTG	TTAAAGCGT	TCATTCTAAG											1498
ATAGAATTAT	TAGAGTT															1515

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

-115-

Met	Lys	Glu	Ser	Phe	Tyr	Ile	Glu	Gly	Met	Thr	Cys	Thr	Ala	Cys	Ser	1	5	10	15
Ser	Gly	Ile	Glu	Arg	Ser	Leu	Gly	Arg	Lys	Ser	Phe	Val	Lys	Lys	Ile	20	25	30	
Glu	Val	Ser	Leu	Leu	Asn	Lys	Ser	Ala	Asn	Ile	Glu	Phe	Asp	Glu	Asn	35	40	45	
Gln	Thr	Asn	Leu	Asp	Glu	Ile	Phe	Lys	Leu	Ile	Glu	Lys	Leu	Gly	Tyr	50	55	60	
Ser	Pro	Lys	Lys	Ala	Leu	Thr	Lys	Glu	Lys	Lys	Glu	Phe	Phe	Ser	Pro	65	70	75	80
Asn	Val	Lys	Leu	Ala	Leu	Ala	Val	Ile	Phe	Thr	Leu	Phe	Val	Val	Tyr	85	90	95	
Leu	Ser	Met	Gly	Ala	Met	Leu	Ser	Pro	Ser	Leu	Leu	Pro	Glu	Ser	Leu	100	105	110	
Leu	Ala	Ile	Asp	Asn	His	Ser	Asn	Phe	Leu	Asn	Ala	Cys	Leu	Gln	Leu	115	120	125	
Ile	Gly	Ala	Leu	Ile	Val	Met	His	Leu	Gly	Arg	Asp	Phe	Tyr	Ile	Gln	130	135	140	
Gly	Phe	Lys	Ala	Leu	Trp	His	Arg	Gln	Pro	Asn	Met	Ser	Ser	Leu	Ile	145	150	155	160
Ala	Ile	Gly	Thr	Ser	Ala	Ala	Leu	Ile	Ser	Ser	Leu	Trp	Gln	Leu	Tyr	165	170	175	
Leu	Val	Tyr	Thr	Asn	His	Tyr	Thr	Asp	Gln	Trp	Ser	Tyr	Gly	His	Tyr	180	185	190	
Tyr	Phe	Glu	Ser	Val	Cys	Val	Ile	Leu	Met	Phe	Val	Met	Val	Gly	Lys	195	200	205	
Arg	Ile	Glu	Asn	Val	Ser	Lys	Asp	Lys	Ala	Leu	Asp	Ala	Met	Gln	Ala	210	215	220	
Leu	Met	Lys	Asn	Ala	Pro	Lys	Thr	Ala	Leu	Lys	Met	Gln	Asn	Asn	Gln	225	230	235	240
Gln	Ile	Glu	Val	Leu	Val	Asp	Ser	Ile	Val	Val	Gly	Asp	Ile	Leu	Lys	245	250	255	
Val	Leu	Pro	Gly	Ser	Ala	Ile	Ala	Val	Asp	Gly	Glu	Ile	Ile	Glu	Gly	260	265	270	
Glu	Gly	Glu	Leu	Asp	Glu	Ser	Met	Leu	Ser	Gly	Glu	Ala	Leu	Pro	Val	275	280	285	
Tyr	Lys	Lys	Val	Gly	Asp	Lys	Val	Phe	Ser	Gly	Thr	Phe	Asn	Ser	His	290	295	300	
Thr	Ser	Phe	Leu	Met	Lys	Ala	Thr	Gln	Asn	Asn	Lys	Asn	Ser	Thr	Leu	305	310	315	320
Ser	Gln	Ile	Ile	Glu	Met	Ile	Tyr	Asn	Ala	Gln	Ser	Ser	Lys	Ala	Glu	325	330	335	
Ile	Ser	Arg	Leu	Ala	Asp	Lys	Val	Ser	Ser	Val	Phe	Val	Pro	Ser	Val	340	345	350	
Ile	Ala	Ile	Ser	Ile	Leu	Ala	Phe	Val	Val	Trp	Leu	Ile	Ile	Ala	Pro	355	360	365	
Lys	Pro	Asp	Phe	Trp	Trp	Asn	Phe	Gly	Ile	Ala	Leu	Glu	Val	Phe	Val	370	375	380	
Ser	Val	Leu	Val	Ile	Ser	Cys	Pro	Cys	Ala	Leu	Gly	Leu	Leu	Arg	Leu	385	390	395	400

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1443 base pairs

-116-

- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 76...1389
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ACTTTAAAAA ACCCCCTTAA AAAGGTTTTT AGGTATAATT AGCGATCTTT TAGTTTCAAA	60
TAGTAGAGAG ATGGG ATG AAA AAA ATA TGG CTT TTA GTG TGG GGC TTG TGT	111
Met Lys Lys Ile Trp Leu Leu Val Trp Gly Leu Cys	
1 5 10	
TCT TGG GTG TTT TTG CAT GCG ATA GAG ATG ATA GAA AAA GCC CCT ACA	159
Ser Trp Val Phe Leu His Ala Ile Glu Met Ile Glu Lys Ala Pro Thr	
15 20 25	
AAT GTA GAG GAT AGA GAC AAA GCC CCC CAT TTG TTG CTT TTA GCA GGG	207
Asn Val Glu Asp Arg Asp Lys Ala Pro His Leu Leu Leu Leu Ala Gly	
30 35 40	
ATT CAA GGC GAT GAG CCT GGT GGG TTT AAT GCA ACT AAT TTG TTT TTA	255
Ile Gln Gly Asp Glu Pro Gly Gly Phe Asn Ala Thr Asn Leu Phe Leu	
45 50 55 60	
ATG CAT TAT AGC GTT TTA AAA GGT TTG GTT GAA GTG GTT CCT GTA TTG	303
Met His Tyr Ser Val Leu Lys Gly Leu Val Glu Val Val Pro Val Leu	
65 70 75	
AAT AAG CCT TCC ATG TTA AGA AAT CAT AGG GGC TTG TAT GGG GAT ATG	351
Asn Lys Pro Ser Met Leu Arg Asn His Arg Gly Leu Tyr Gly Asp Met	
80 85 90	
AAC CGC AAA TTT GCC GCT TTA GAC AAG AAT GAC CCT GAA TAC CCC ACT	399
Asn Arg Lys Phe Ala Ala Leu Asp Lys Asn Asp Pro Glu Tyr Pro Thr	
95 100 105	
ATC CAG GAA ATC AAA TCC TTG ATT GCA AAA CCC AGT ATA GAC GCT GTC	447
Ile Gln Glu Ile Lys Ser Leu Ile Ala Lys Pro Ser Ile Asp Ala Val	
110 115 120	
TTG CAT TTG CAT GAT GGC GGT GGG TAT TAC CGC CCT GTT TAT GTT GAT	495
Leu His Leu His Asp Gly Gly Tyr Tyr Arg Pro Val Tyr Val Asp	
125 130 135 140	
GCG ATG CTC AAT CCT AAG CGC TGG GGG AAT TGC TTT ATT ATT GAT CAA	543
Ala Met Leu Asn Pro Lys Arg Trp Gly Asn Cys Phe Ile Ile Asp Gln	
145 150 155	

-117-

GAT	GAG	GTT	AAA	GGG	GCG	AAA	TTC	CCT	AAT	TTG	CTT	GCT	TTT	GCA	AAC	591
Asp	Glu	Val	Lys	Gly	Ala	Lys	Phe	Pro	Asn	Leu	Leu	Ala	Phe	Ala	Asn	
			160					165					170			
AAT	ACG	ATT	GAG	AGT	ATC	AAC	GCC	CAT	TTA	TTG	CAC	CCC	ATT	GAA	GAG	639
Asn	Thr	Ile	Glu	Ser	Ile	Asn	Ala	His	Leu	Leu	His	Pro	Ile	Glu	Glu	
		175					180					185				
TAT	CAT	TTA	AAA	AAC	ACG	CGC	ACC	GCG	CAA	GGC	GAT	ACA	GAA	ATG	CAA	687
Tyr	His	Leu	Lys	Asn	Thr	Arg	Thr	Ala	Gln	Gly	Asp	Thr	Glu	Met	Gln	
	190					195					200					
AAA	GCC	CTA	ACT	TTT	TAT	GCG	ATC	AAC	CAA	AAA	AAG	AGC	GCT	TTT	GCC	735
Lys	Ala	Leu	Thr	Phe	Tyr	Ala	Ile	Asn	Gln	Lys	Lys	Ser	Ala	Phe	Ala	
205					210				215					220		
AAT	GAA	GCT	AGC	AAA	GAA	CTC	CCT	TTA	GCA	TCA	AGG	GTG	TTT	TAC	CAC	783
Asn	Glu	Ala	Ser	Lys	Glu	Leu	Pro	Leu	Ala	Ser	Arg	Val	Phe	Tyr	His	
			225					230						235		
CTG	CAA	GCC	ATT	GAG	GGC	TTA	CTC	AAT	CAG	CTC	AAT	ATC	CCT	TTT	AAG	831
Leu	Gln	Ala	Ile	Glu	Gly	Leu	Leu	Asn	Gln	Leu	Asn	Ile	Pro	Phe	Lys	
			240					245					250			
CGC	GAT	TTT	GAT	CTT	AAC	CCT	AAC	AGC	GTG	CAT	GCC	CTA	ATC	AAT	GAT	879
Arg	Asp	Phe	Asp	Leu	Asn	Pro	Asn	Ser	Val	His	Ala	Leu	Ile	Asn	Asp	
		255					260					265				
AAA	AAC	TTG	TGG	GCA	AAA	ATC	AGC	TCT	TTG	CCT	AAA	ATG	CCC	CTT	TTT	927
Lys	Asn	Leu	Trp	Ala	Lys	Ile	Ser	Ser	Leu	Pro	Lys	Met	Pro	Leu	Phe	
	270					275					280					
AAC	TTG	CGC	CCT	AAA	CTC	AAT	CAT	TTC	CCC	TTA	CCC	CAC	AAC	ACT	AAA	975
Asn	Leu	Arg	Pro	Lys	Leu	Asn	His	Phe	Pro	Leu	Pro	His	Asn	Thr	Lys	
285					290					295				300		
ATC	CCA	CAA	ATC	CCC	ATA	GAG	AGC	AAC	GCT	TAC	ATT	GTA	GGG	CTA	GTC	1023
Ile	Pro	Gln	Ile	Pro	Ile	Glu	Ser	Asn	Ala	Tyr	Ile	Val	Gly	Leu	Val	
				305					310					315		
AAA	AAT	AAA	CAA	GAA	GTG	TTT	TTA	AAA	TAC	GGC	AAC	AAG	CTC	ATG	ACA	1071
Lys	Asn	Lys	Gln	Glu	Val	Phe	Leu	Lys	Tyr	Gly	Asn	Lys	Leu	Met	Thr	
			320					325					330			
CGA	TTA	TCG	CCT	TTT	TAC	ATA	GAG	TTT	GAT	CCT	TCT	TTA	GAA	GAA	GTG	1119
Arg	Leu	Ser	Pro	Phe	Tyr	Ile	Glu	Phe	Asp	Pro	Ser	Leu	Glu	Glu	Val	
		335					340					345				
AAA	ATG	CAA	ATT	GAC	AAT	AAG	GAT	CAA	ATG	GTT	AAA	ATA	GGG	AGC	GTG	1167
Lys	Met	Gln	Ile	Asp	Asn	Lys	Asp	Gln	Met	Val	Lys	Ile	Gly	Ser	Val	
	350					355					360					
GTT	GAA	GTG	AAA	GAG	AGT	TTT	TAT	ATC	CAT	GCT	ATG	GAC	AAT	ATC	CGT	1215
Val	Glu	Val	Lys	Glu	Ser	Phe	Tyr	Ile	His	Ala	Met	Asp	Asn	Ile	Arg	
365					370				375					380		

-118-

GCG AAT GTG ATT GGC TTT AGC GTT TCT AAT GAA AAT AAG CCT AAT GAA 1263
 Ala Asn Val Ile Gly Phe Ser Val Ser Asn Glu Asn Lys Pro Asn Glu
 385 390 395

GCG GGT TAT ACG ATT AAA TTT AAA GAT TTT CAA AAA CGC TTT TCA TTG 1311
 Ala Gly Tyr Thr Ile Lys Phe Lys Asp Phe Gln Lys Arg Phe Ser Leu
 400 405 410

GAC AAG CAA GAA AGG ATC TAT CGC ATA GAA TTT TAT AAA AAC AAC GCG 1359
 Asp Lys Gln Glu Arg Ile Tyr Arg Ile Glu Phe Tyr Lys Asn Asn Ala
 415 420 425

TTT AGC GGG ATG ATC TTA GTG AAA TTT GTG TAGGAATGGA TAAATCTCAT TGC 1412
 Phe Ser Gly Met Ile Leu Val Lys Phe Val
 430 435

CTTTTAACAT TCAAGGGTTT TGGTATTTTT T 1443

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 438 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Lys Lys Ile Trp Leu Leu Val Trp Gly Leu Cys Ser Trp Val Phe
 1 5 10 15
 Leu His Ala Ile Glu Met Ile Glu Lys Ala Pro Thr Asn Val Glu Asp
 20 25 30
 Arg Asp Lys Ala Pro His Leu Leu Leu Ala Gly Ile Gln Gly Asp
 35 40 45
 Glu Pro Gly Gly Phe Asn Ala Thr Asn Leu Phe Leu Met His Tyr Ser
 50 55 60
 Val Leu Lys Gly Leu Val Glu Val Val Pro Val Leu Asn Lys Pro Ser
 65 70 75 80
 Met Leu Arg Asn His Arg Gly Leu Tyr Gly Asp Met Asn Arg Lys Phe
 85 90 95
 Ala Ala Leu Asp Lys Asn Asp Pro Glu Tyr Pro Thr Ile Gln Glu Ile
 100 105 110
 Lys Ser Leu Ile Ala Lys Pro Ser Ile Asp Ala Val Leu His Leu His
 115 120 125
 Asp Gly Gly Gly Tyr Tyr Arg Pro Val Tyr Val Asp Ala Met Leu Asn
 130 135 140
 Pro Lys Arg Trp Gly Asn Cys Phe Ile Ile Asp Gln Asp Glu Val Lys
 145 150 155 160
 Gly Ala Lys Phe Pro Asn Leu Leu Ala Phe Ala Asn Asn Thr Ile Glu
 165 170 175
 Ser Ile Asn Ala His Leu Leu His Pro Ile Glu Glu Tyr His Leu Lys
 180 185 190

-119-

```

Asn Thr Arg Thr Ala Gln Gly Asp Thr Glu Met Gln Lys Ala Leu Thr
    195                200                205
Phe Tyr Ala Ile Asn Gln Lys Lys Ser Ala Phe Ala Asn Glu Ala Ser
    210                215                220
Lys Glu Leu Pro Leu Ala Ser Arg Val Phe Tyr His Leu Gln Ala Ile
    225                230                235                240
Glu Gly Leu Leu Asn Gln Leu Asn Ile Pro Phe Lys Arg Asp Phe Asp
                245                250                255
Leu Asn Pro Asn Ser Val His Ala Leu Ile Asn Asp Lys Asn Leu Trp
                260                265                270
Ala Lys Ile Ser Ser Leu Pro Lys Met Pro Leu Phe Asn Leu Arg Pro
    275                280                285
Lys Leu Asn His Phe Pro Leu Pro His Asn Thr Lys Ile Pro Gln Ile
    290                295                300
Pro Ile Glu Ser Asn Ala Tyr Ile Val Gly Leu Val Lys Asn Lys Gln
    305                310                315                320
Glu Val Phe Leu Lys Tyr Gly Asn Lys Leu Met Thr Arg Leu Ser Pro
                325                330                335
Phe Tyr Ile Glu Phe Asp Pro Ser Leu Glu Glu Val Lys Met Gln Ile
                340                345                350
Asp Asn Lys Asp Gln Met Val Lys Ile Gly Ser Val Val Glu Val Lys
    355                360                365
Glu Ser Phe Tyr Ile His Ala Met Asp Asn Ile Arg Ala Asn Val Ile
    370                375                380
Gly Phe Ser Val Ser Asn Glu Asn Lys Pro Asn Glu Ala Gly Tyr Thr
    385                390                395                400
Ile Lys Phe Lys Asp Phe Gln Lys Arg Phe Ser Leu Asp Lys Gln Glu
                405                410                415
Arg Ile Tyr Arg Ile Glu Phe Tyr Lys Asn Asn Ala Phe Ser Gly Met
                420                425                430
Ile Leu Val Lys Phe Val
    435

```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1280 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 66...1223
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

ATCAATACCC CTTAAATAAA AGATATAATG CTGTATTATA AGCTAGTTTT AATTACAATT      60
TTCAA ATG TTA AGG AAA AAC ATT TTA GCT TAC TAT GGG GCG AAT TTT CTC      110
Met Leu Arg Lys Asn Ile Leu Ala Tyr Tyr Gly Ala Asn Phe Leu
    1             5             10             15

```


-120-

TTA ATC ATC GCT CAA AGC TTA CCC CAT GCG ATT TTA ACC CCC TTG TTG	158
Leu Ile Ile Ala Gln Ser Leu Pro His Ala Ile Leu Thr Pro Leu Leu	
20 25 30	
CTT TCT AAA GGG CTT AGT TTG AGT GAA ATC TTG CTC GTG CAA ACC TTT	206
Leu Ser Lys Gly Leu Ser Leu Ser Glu Ile Leu Leu Val Gln Thr Phe	
35 40 45	
TTT AGC TTT TGC GTG CTA GTG GCT GAA TAC CCA AGC GGC GTT TTA GCG	254
Phe Ser Phe Cys Val Leu Val Ala Glu Tyr Pro Ser Gly Val Leu Ala	
50 55 60	
GAT TTG ATG AGC CGA AAA AAT TTA TTC CTG GTT TCT AAT GCC TTT TTA	302
Asp Leu Met Ser Arg Lys Asn Leu Phe Leu Val Ser Asn Ala Phe Leu	
65 70 75	
ATC GCT AGT TTT TCG TTT GTG CTG TTT TTT GAT AGC TTT ATT TTC ATG	350
Ile Ala Ser Phe Ser Phe Val Leu Phe Phe Asp Ser Phe Ile Phe Met	
80 85 90 95	
CTT TTA GCG TGG GGG TTG TAT GGT TTG TAT AGC GCA TGC TCT AGC GGC	398
Leu Leu Ala Trp Gly Leu Tyr Gly Leu Tyr Ser Ala Cys Ser Ser Gly	
100 105 110	
ACG ATT GAA GCT TCA CTC ATC ACA GAC ATT AAG GAA AAC AAA AAA GAT	446
Thr Ile Glu Ala Ser Leu Ile Thr Asp Ile Lys Glu Asn Lys Lys Asp	
115 120 125	
TTA TCC AAG TTT TTA GCC AAA AAC AAT CAA ATT ACT TAT TTA GGC ATG	494
Leu Ser Lys Phe Leu Ala Lys Asn Asn Gln Ile Thr Tyr Leu Gly Met	
130 135 140	
ATT ATA GGG AGT TCT TTG GGA TCG TTT TTG TAT CTC AAA GTC CAT GCG	542
Ile Ile Gly Ser Ser Leu Gly Ser Phe Leu Tyr Leu Lys Val His Ala	
145 150 155	
ATG CTG TAT ATT GTG GGG ATT TTT TTA ATC ATG CTC TGT GTG CTA ACG	590
Met Leu Tyr Ile Val Gly Ile Phe Leu Ile Met Leu Cys Val Leu Thr	
160 165 170 175	
ATC ATT TTT TAT TTT AAA GAG AAA GAA GGG GAT TTT AAA AGC CAA AAA	638
Ile Ile Phe Tyr Phe Lys Glu Lys Glu Gly Asp Phe Lys Ser Gln Lys	
180 185 190	
AGC CTG AAA CTC CTT AAA GAG CAA GTC AAA GGC AGT CTT AAA GAG CTT	686
Ser Leu Lys Leu Leu Lys Glu Gln Val Lys Gly Ser Leu Lys Glu Leu	
195 200 205	
AAA GAT AAC CCC AAA CTT AAA ATT CTG TTA GTG GGG CAT TTG ATT ACG	734
Lys Asp Asn Pro Lys Leu Lys Ile Leu Leu Val Gly His Leu Ile Thr	
210 215 220	
CCC GTC TTT TTT ATG AGC CAT TTT CAA ATG TGG CAA GCG TAT TTT TTA	782
Pro Val Phe Phe Met Ser His Phe Gln Met Trp Gln Ala Tyr Phe Leu	
225 230 235	

-121-

AAA CAA GGC GTT AAA GAG CAA TAC CTT TTT GTG TTT TAT ATC GCT TTT	830
Lys Gln Gly Val Lys Glu Gln Tyr Leu Phe Val Phe Tyr Ile Ala Phe	
240 245 250 255	
CAA GTG ATT TCT ATT CTC ATT CAT TTT TTA AAA GCC TCT AGT TAT AGC	878
Gln Val Ile Ser Ile Leu Ile His Phe Leu Lys Ala Ser Ser Tyr Ser	
260 265 270	
CAA AAA ATC GCC TTG AGT TCG CTT GTG GTG TTG TTA GGC GTT AGC CCC	926
Gln Lys Ile Ala Leu Ser Ser Leu Val Val Leu Leu Gly Val Ser Pro	
275 280 285	
TTA TTG CTT AGC AAT ATC CCT TAT TGT TTC ATA GGG GTG TAT GCG CTC	974
Leu Leu Leu Ser Asn Ile Pro Tyr Cys Phe Ile Gly Val Tyr Ala Leu	
290 295 300	
ATG GTG GCG TTT TTC ACT TAC ATG AGC TAT TGC TTA AAC TAT CAA TTC	1022
Met Val Ala Phe Phe Thr Tyr Met Ser Tyr Cys Leu Asn Tyr Gln Phe	
305 310 315	
TCC AAA TTC GTT TCT AAA AAC AAC ATT TCC TCG CTC TCA TCG CTT TTA	1070
Ser Lys Phe Val Ser Lys Asn Asn Ile Ser Ser Leu Ser Ser Leu Leu	
320 325 330 335	
TCA AGC TGT GTG CGC GTG GTC TCT GTG CTA ATC TTA TCG CTC AGC AGT	1118
Ser Ser Cys Val Arg Val Val Ser Val Leu Ile Leu Ser Leu Ser Ser	
340 345 350	
CTG GAA CTG CGT TAC TTC TCA CCC CTA ACT ATC ATA ACC ATG CAT TTT	1166
Leu Glu Leu Arg Tyr Phe Ser Pro Leu Thr Ile Ile Thr Met His Phe	
355 360 365	
GCC TTG ACG CTT ATC ATC CTC TTT TTC TTT TTG TAT AAG GCT AAG CCG	1214
Ala Leu Thr Leu Ile Ile Leu Phe Phe Phe Leu Tyr Lys Ala Lys Pro	
370 375 380	
TTT GAT GAG TGAGCGGCTT TAAGAGTGCA ACCTTTTAGC GATTTCTATA GCAACATCA	1272
Phe Asp Glu	
385	
TAGCCATG	1280

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 386 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Leu Arg Lys Asn Ile Leu Ala Tyr Tyr Gly Ala Asn Phe Leu Leu

-122-

1	5	10	15
Ile Ile Ala Gln Ser Leu Pro His Ala Ile Leu Thr Pro Leu Leu Leu			
20	25	30	
Ser Lys Gly Leu Ser Leu Ser Glu Ile Leu Leu Val Gln Thr Phe Phe			
35	40	45	
Ser Phe Cys Val Leu Val Ala Glu Tyr Pro Ser Gly Val Leu Ala Asp			
50	55	60	
Leu Met Ser Arg Lys Asn Leu Phe Leu Val Ser Asn Ala Phe Leu Ile			
65	70	75	80
Ala Ser Phe Ser Phe Val Leu Phe Phe Asp Ser Phe Ile Phe Met Leu			
85	90	95	
Leu Ala Trp Gly Leu Tyr Gly Leu Tyr Ser Ala Cys Ser Ser Gly Thr			
100	105	110	
Ile Glu Ala Ser Leu Ile Thr Asp Ile Lys Glu Asn Lys Lys Asp Leu			
115	120	125	
Ser Lys Phe Leu Ala Lys Asn Asn Gln Ile Thr Tyr Leu Gly Met Ile			
130	135	140	
Ile Gly Ser Ser Leu Gly Ser Phe Leu Tyr Leu Lys Val His Ala Met			
145	150	155	160
Leu Tyr Ile Val Gly Ile Phe Leu Ile Met Leu Cys Val Leu Thr Ile			
165	170	175	
Ile Phe Tyr Phe Lys Glu Lys Glu Gly Asp Phe Lys Ser Gln Lys Ser			
180	185	190	
Leu Lys Leu Leu Lys Glu Gln Val Lys Gly Ser Leu Lys Glu Leu Lys			
195	200	205	
Asp Asn Pro Lys Leu Lys Ile Leu Leu Val Gly His Leu Ile Thr Pro			
210	215	220	
Val Phe Phe Met Ser His Phe Gln Met Trp Gln Ala Tyr Phe Leu Lys			
225	230	235	240
Gln Gly Val Lys Glu Gln Tyr Leu Phe Val Phe Tyr Ile Ala Phe Gln			
245	250	255	
Val Ile Ser Ile Leu Ile His Phe Leu Lys Ala Ser Ser Tyr Ser Gln			
260	265	270	
Lys Ile Ala Leu Ser Ser Leu Val Val Leu Leu Gly Val Ser Pro Leu			
275	280	285	
Leu Leu Ser Asn Ile Pro Tyr Cys Phe Ile Gly Val Tyr Ala Leu Met			
290	295	300	
Val Ala Phe Phe Thr Tyr Met Ser Tyr Cys Leu Asn Tyr Gln Phe Ser			
305	310	315	320
Lys Phe Val Ser Lys Asn Asn Ile Ser Ser Leu Ser Ser Leu Leu Ser			
325	330	335	
Ser Cys Val Arg Val Val Ser Val Leu Ile Leu Ser Leu Ser Ser Leu			
340	345	350	
Glu Leu Arg Tyr Phe Ser Pro Leu Thr Ile Ile Thr Met His Phe Ala			
355	360	365	
Leu Thr Leu Ile Ile Leu Phe Phe Leu Tyr Lys Ala Lys Pro Phe			
370	375	380	
Asp Glu			
385			

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1264 base pairs

(B) TYPE: nucleic acid

-123-

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 51...1205

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATTAAATATG ACTATATACA CTACAACAAT AAGATTTTGA AAGGTTGGTA	ATG GAA	56
	Met Glu	
	1	
TCA GTA AAA ACA GGA AAA ACA AAT AAG GTT GGC AAG AAT ACA GAG ATG		104
Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr Glu Met		
5 10 15		
GCT AAT ACA AAG GCA AAT AAA GAG GCT CAT TTT AAA CAA GCG AGC ACC		152
Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala Ser Thr		
20 25 30		
ATT ACA AAT ATA ATC AGA TCA ATT CGT GGG ATT TTT ACA AAA ATT GCA		200
Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys Ile Ala		
35 40 45 50		
AAG AAA GTT AGA GGA CTT GTA AAA AAA CAC CCC AAG AAA AGC AGT GCG		248
Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser Ser Ala		
55 60 65		
GCA TTA GTA GTA TTG ACC CAT ATT GCG TGC AAG AAA GCG AAA GAA TTA		296
Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys Glu Leu		
70 75 80		
GAC GAT AAA GTC CAA GAT AAA TCC AAA CAA GCT GAA AAA GAA AAT CAA		344
Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu Asn Gln		
85 90 95		
ATC AAT TGG TGG AAA TAT TCA GGA TTA ACA ATA GCG ACA AGT TTA TTA		392
Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser Leu Leu		
100 105 110		
TTA GCC GCT TGT AGC ACT GGT GAT GTT AGT GAA CAA ATA GAA CTA GAA		440
Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln Ile Glu Leu Glu		
115 120 125 130		
CAA GAA AAA CAA AAG ACG AGC AAT ATA GAG ACT AAC AAT CAA ATA AAA		488
Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln Ile Lys		
135 140 145		
GTA GAA CAA GAA AAA CAA AAG ACA AGC AAT ATA GAG ACT AAT AAT CAA		536
Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln		
150 155 160		

-124-

ATA AAA GTA GAA CAA GAA CAA CAG AAA ACA GAA CAA GAA MGA CAG AAA	584
Ile Lys Val Glu Gln Glu Gln Gln Lys Thr Glu Gln Glu Xaa Gln Lys	
165 170 175	
ACA GAA CAA GAA AGA CAG AAG ACA GAA CAA GAA AAA CAA AAG ACC ATT	632
Thr Glu Gln Glu Arg Gln Lys Thr Glu Gln Glu Lys Gln Lys Thr Ile	
180 185 190	
AAA ACA CAG AAA GAT TTC ATT AAA TAT GTA GAA CAA AAT TGC CAA GAA	680
Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys Gln Glu	
195 200 205 210	
AAT CAT AAT CAA TTC TTT ATT GAA AAA GGA GGA ATT AAG GCT GGT ATT	728
Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala Gly Ile	
215 220 225	
GGT ATA GAA GTA GAA GCT GAA TGC AAA ACC CCT AAA CCT GCA AAA ACC	776
Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr	
230 235 240	
AAT CAA ACC CCT ATC CAG CCA AAA CAC CTC CCA AAC TCT AAA CAA CCC	824
Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln Pro	
245 250 255	
CGC TCT CAA AGA GGA TCA AAA GCG CAA GAG CTT ATC GCT TAT TTG CAA	872
Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr Leu Gln	
260 265 270	
AAA GAG CTA GAA TCT CTG CCC TAT TCA CAA AAA GCT ATC GCT AAA CAA	920
Lys Glu Leu Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys Gln	
275 280 285 290	
GTG GAT TTT TAT AGA CCA AGT TCT ATC GCT TAT TTA GAA CTA GAC CCT	968
Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu Asp Pro	
295 300 305	
AGA GAT TTT AAT GTT ACA GAA GAA TGG CAA AAA GAA AAT TTA AAA ATA	1016
Arg Asp Phe Asn Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys Ile	
310 315 320	
CGC TCT AAA GCT CAA GCT AAA ATG CTT GAA ATG AGG AGT TTA AAA CCA	1064
Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Ser Leu Lys Pro	
325 330 335	
GAC TCA CAA GCC CAC CTT TCA ACC TCT CAA AGC CTT TTG TTC GTT CAA	1112
Asp Ser Gln Ala His Leu Ser Thr Ser Gln Ser Leu Leu Phe Val Gln	
340 345 350	
AAA ATA TTT GCT GAT GTT AAT AAA GAA ATA AAA GTA GTT GCT AAT ACT	1160
Lys Ile Phe Ala Asp Val Asn Lys Glu Ile Lys Val Val Ala Asn Thr	
355 360 365 370	
GAA AAG AAA GCA GAA AAA GCG GGT TAT GGT TAT AGT AAA AGG ATG TAGGC	1210
Glu Lys Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met	
375 380 385	

-125-

ATAAGAAAAC ACCATAAAAT CGTTCCTTAGC TTATTTATAG TATTTTAAAA ACTC

1264

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 385 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```

Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr
 1           5           10           15
Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala
          20           25           30
Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys
          35           40           45
Ile Ala Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser
          50           55           60
Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys
          65           70           75           80
Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu
          85           90           95
Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser
          100          105          110
Leu Leu Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln Ile Glu
          115          120          125
Leu Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln
          130          135          140
Ile Lys Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn
          145          150          155          160
Asn Gln Ile Lys Val Glu Gln Glu Gln Gln Lys Thr Glu Gln Glu Xaa
          165          170          175
Gln Lys Thr Glu Gln Glu Arg Gln Lys Thr Glu Gln Glu Lys Gln Lys
          180          185          190
Thr Ile Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys
          195          200          205
Gln Glu Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala
          210          215          220
Gly Ile Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala
          225          230          235          240
Lys Thr Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys
          245          250          255
Gln Pro Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr
          260          265          270
Leu Gln Lys Glu Leu Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala
          275          280          285
Lys Gln Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu
          290          295          300
Asp Pro Arg Asp Phe Asn Val Thr Glu Glu Trp Gln Lys Glu Asn Leu
          305          310          315          320

```

[illegible]

(A) LENGTH: 410 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 62...340
(D) OTHER INFORMATION:

ATTCATTTTAC	TTTTGAGAAA	TATAATTCTC	TCGTTTAA	GATCATACA	AGGAGTTTCG	60
T ATG AAA AAG CAA ATC TTG ACA GGT GTT TTA TTA TCA GTT TTG GCA GTG	109					
Met Lys Lys Gln Ile Leu Thr Gly Val Leu Leu Ser Val Leu Ala Val						
1	5	10	15			
AGT TCT GCA TAC GCT CAC AAA GAT AAA AAA GAC GCC AAA AAA CCT AAA	157					
Ser Ser Ala Tyr Ala His Lys Asp Lys Lys Asp Ala Lys Lys Pro Lys						
20	25	30				
TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC GAC AAA AAA GAC GCT AAA	205					
Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys Asp Ala Lys						
35	40	45				
AAA CCT AAA TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC GAC AAA AAA	253					
Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys						
50	55	60				
GAC GCT AAA AAA CCT AAA TTT AGC ACA GAA TTA GTC GTG GCT CAA AAC	301					
Asp Ala Lys Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn						
65	70	75	80			
GAC AAA AAA GAC GCT AAA AAA CCT AAA AAC TCA GTG GTC TAATGGCTTT GA	352					
Asp Lys Lys Asp Ala Lys Lys Pro Lys Asn Ser Val Val						
85	90					
CTCTAAAAAA	GCGTTTTTAA	AAACGCTTTT	TTGGATATTA	TCCTATAATT	TCCTACCA	410

-127-

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

Met Lys Lys Gln Ile Leu Thr Gly Val Leu Leu Ser Val Leu Ala Val
 1             5             10             15
Ser Ser Ala Tyr Ala His Lys Asp Lys Lys Asp Ala Lys Lys Pro Lys
      20             25             30
Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys Asp Ala Lys
      35             40             45
Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn Asp Lys Lys
      50             55             60
Asp Ala Lys Lys Pro Lys Phe Ser Thr Glu Leu Val Val Ala Gln Asn
65             70             75             80
Asp Lys Lys Asp Ala Lys Lys Pro Lys Asn Ser Val Val
      85             90

```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2097 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 67...2046
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```

TAAAAACCCC TATCATAGGG CGTGGCATGA AGAAAAAAGC AAAAGTCTTT TGGTATTGTT      60
TTAATC ATG ATT TAT TGG TTG TAT TTG GCG GTC TTT TTT TTG TTG AGC      108
    Met Ile Tyr Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser
      1             5             10

GCA TTA GAC GCT AAA GAA ATC GCT ATG CAA CGA TTT GAC AAA CAA AAC      156
Ala Leu Asp Ala Lys Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn
15             20             25             30

CAT AAG ATT TTT GAA ATC CTT GCG GAT AAA GTG AGC GCT AAA GAC AAT      204
His Lys Ile Phe Glu Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn

```


-128-

GTG	ATA	ACC	GCA	TCA	GGG	AAT	GCG	ATC	TTA	TTG	AAT	TAT	GAT	GTG	TAT	252
Val	Ile	Thr	Ala	Ser	Gly	Asn	Ala	Ile	Leu	Leu	Asn	Tyr	Asp	Val	Tyr	
			50					55					60			
ATT	CTA	GCG	GAC	AAG	GTG	CGT	TAT	GAC	ACT	AAA	ACC	AAA	GAA	GCG	TTA	300
Ile	Leu	Ala	Asp	Lys	Val	Arg	Tyr	Asp	Thr	Lys	Thr	Lys	Glu	Ala	Leu	
		65					70					75				
TTA	GAG	GGG	AAT	ATC	AAG	GTT	TAT	AGG	GGC	GAG	GGT	TTG	CTC	GTT	AAA	348
Leu	Glu	Gly	Asn	Ile	Lys	Val	Tyr	Arg	Gly	Glu	Gly	Leu	Leu	Val	Lys	
	80					85					90					
ACC	GAT	TAC	GTG	AAA	TTG	AGT	TTG	AAT	GAA	AAA	TAT	GAA	ATC	ATT	TTC	396
Thr	Asp	Tyr	Val	Lys	Leu	Ser	Leu	Asn	Glu	Lys	Tyr	Glu	Ile	Ile	Phe	
	95				100					105					110	
CCC	TTT	TAT	GTC	CAA	GAC	AGC	GTG	AGC	GGG	ATT	TGG	GTG	AGC	GCG	GAT	444
Pro	Phe	Tyr	Val	Gln	Asp	Ser	Val	Ser	Gly	Ile	Trp	Val	Ser	Ala	Asp	
				115					120					125		
ATT	GCC	AGC	GGA	AAG	GAT	CAA	AAA	TAT	AAG	GTT	AAA	AAC	ATG	AGC	ACT	492
Ile	Ala	Ser	Gly	Lys	Asp	Gln	Lys	Tyr	Lys	Val	Lys	Asn	Met	Ser	Thr	
			130					135					140			
TCA	GGG	TGC	AGC	ATT	GAT	AAC	CCC	ATT	TGG	CAT	GTC	AAT	GCG	ACT	TCA	540
Ser	Gly	Cys	Ser	Ile	Asp	Asn	Pro	Ile	Trp	His	Val	Asn	Ala	Thr	Ser	
		145					150						155			
GGC	TCA	TTC	AAC	ATG	CAA	AAA	TCG	CAT	TTG	TCT	ATG	TGG	AAT	CCT	AAG	588
Gly	Ser	Phe	Asn	Met	Gln	Lys	Ser	His	Leu	Ser	Met	Trp	Asn	Pro	Lys	
	160					165					170					
ATC	TAT	GTC	GGT	GAT	ATT	CCT	GTA	TTG	TAT	TTG	CCC	TAT	ATT	TTC	ATG	636
Ile	Tyr	Val	Gly	Asp	Ile	Pro	Val	Leu	Tyr	Leu	Pro	Tyr	Ile	Phe	Met	
	175				180					185					190	
TCC	ACG	AGC	AAT	AAA	AGA	ACT	ACT	GGG	TTT	TTA	TAC	CCT	GAG	TTT	GGC	684
Ser	Thr	Ser	Asn	Lys	Arg	Thr	Thr	Gly	Phe	Leu	Tyr	Pro	Glu	Phe	Gly	
				195				200						205		
ACT	TCC	AAC	TTA	GAC	GGC	TTT	ATT	TAT	TTG	CAA	CCC	TTT	TAT	TTA	GCC	732
Thr	Ser	Asn	Leu	Asp	Gly	Phe	Ile	Tyr	Leu	Gln	Pro	Phe	Tyr	Leu	Ala	
			210					215					220			
CCC	AAA	AAC	TCA	TGG	GAT	ATG	ACC	TTT	ACC	CCA	CAA	ATC	CGC	TAT	AAA	780
Pro	Lys	Asn	Ser	Trp	Asp	Met	Thr	Phe	Thr	Pro	Gln	Ile	Arg	Tyr	Lys	
		225					230					235				
AGG	GGT	TTT	GGC	TTG	AAT	TTT	GAA	GCG	CGC	TAC	ATT	AAC	TCT	AAA	AAC	828
Arg	Gly	Phe	Gly	Leu	Asn	Phe	Glu	Ala	Arg	Tyr	Ile	Asn	Ser	Lys	Asn	
	240					245					250					
GAC	AGG	TTT	TTA	TTC	AAC	GCG	CGC	TAT	TTT	AGG	AAT	TAC	ACC	CAA	TAT	876
Asp	Arg	Phe	Leu	Phe	Asn	Ala	Arg	Tyr	Phe	Arg	Asn	Tyr	Thr	Gln	Tyr	

-129-

255	260	265	270	
GTC AAA CGC TAC GAT TTG AGG AAT CAA AAT ATC TAC GGG TTT GAA TTT				924
Val Lys Arg Tyr Asp Leu Arg Asn Gln Asn Ile Tyr Gly Phe Glu Phe	275	280	285	
TTA AGC TCT AGC AGG GAC ACT TTA CAA AAA TAC TTC CAC CTT AAG TCT				972
Leu Ser Ser Ser Arg Asp Thr Leu Gln Lys Tyr Phe His Leu Lys Ser	290	295	300	
AAT ATT GAC AAC GGG CAT TAC ATT GAC TTT TTA TAC ATG AAC GAT TTG				1020
Asn Ile Asp Asn Gly His Tyr Ile Asp Phe Leu Tyr Met Asn Asp Leu	305	310	315	
GAC TAT GTG CGT TTT GAA AAG GTT AAT AAG CGT ATC ACA GAC GCC ACG				1068
Asp Tyr Val Arg Phe Glu Lys Val Asn Lys Arg Ile Thr Asp Ala Thr	320	325	330	
CAC ATG TCT AGG GCG AAT TAC TAT TTG CAA ACA GAA AAC AAT TAT TAC				1116
His Met Ser Arg Ala Asn Tyr Tyr Leu Gln Thr Glu Asn Asn Tyr Tyr	335	340	345	350
GGC TTG AAT ATC AAG TAT TTT TTA AAC CTG AAT AAA ATC AAC AAT AAC				1164
Gly Leu Asn Ile Lys Tyr Phe Leu Asn Leu Asn Lys Ile Asn Asn Asn	355	360	365	
CGC ACT TTC CAA TCT GTC CCT AAT TTG CAA TAC CAT AAA TAT TTA AAT				1212
Arg Thr Phe Gln Ser Val Pro Asn Leu Gln Tyr His Lys Tyr Leu Asn	370	375	380	
TCT TTG TAT TTT AGA AAT TTG TTG TAT TCG GTG GAT TAT CAG TTT AGA				1260
Ser Leu Tyr Phe Arg Asn Leu Leu Tyr Ser Val Asp Tyr Gln Phe Arg	385	390	395	
AAC ACC GCA AGA GAG ATT GGT TAT GGC TAT GTG CAA AAC GCT TTG AAT				1308
Asn Thr Ala Arg Glu Ile Gly Tyr Gly Tyr Val Gln Asn Ala Leu Asn	400	405	410	
GTG CCG GTG GGC TTG CAA TTT TCT TTG TTT AAA AAG TAT TTG TCT TTA				1356
Val Pro Val Gly Leu Gln Phe Ser Leu Phe Lys Lys Tyr Leu Ser Leu	415	420	425	430
GGG CTT TGG AAT GAT CTC CAA CTA TCT AAT GTG GCT TTA ATG CAA TCT				1404
Gly Leu Trp Asn Asp Leu Gln Leu Ser Asn Val Ala Leu Met Gln Ser	435	440	445	
AAA AAT TCC TTC GTG CCT ACG ATC CCT AAT GAA TCA AGG GAA TTT GGG				1452
Lys Asn Ser Phe Val Pro Thr Ile Pro Asn Glu Ser Arg Glu Phe Gly	450	455	460	
AAT TTT GTG TCT TCA AAT TTT TCC ATG TAT GTC AAT ACG GAT TTG GCT				1500
Asn Phe Val Ser Ser Asn Phe Ser Met Tyr Val Asn Thr Asp Leu Ala	465	470	475	
AGA GAA TAC AAC AAG CTT TTC CAC ACG ATC CAA CTA GAA GCG ATT TTC				1548

-130-

Arg	Glu	Tyr	Asn	Lys	Leu	Phe	His	Thr	Ile	Gln	Leu	Glu	Ala	Ile	Phe		
480						485					490						
AAC	ATC	CCT	TAT	TAC	ACC	TTT	AAA	AAC	GGC	TTA	TTT	TCT	CAA	AAC	ATG	1596	
Asn	Ile	Pro	Tyr	Tyr	Thr	Phe	Lys	Asn	Gly	Leu	Phe	Ser	Gln	Asn	Met		
495					500					505					510		
TAT	GCT	TTA	AGC	GCG	CAA	GCC	TTA	AAC	AGC	TAC	ACT	TCG	CCT	TTA	TTG	1644	
Tyr	Ala	Leu	Ser	Ala	Gln	Ala	Leu	Asn	Ser	Tyr	Thr	Ser	Pro	Leu	Leu		
				515					520					525			
AGA	GAT	TAT	GAT	TAT	CAA	GGG	CGT	TTG	TAT	GAC	TCG	GTG	TGG	AAT	CCT	1692	
Arg	Asp	Tyr	Asp	Tyr	Gln	Gly	Arg	Leu	Tyr	Asp	Ser	Val	Trp	Asn	Pro		
			530					535					540				
AGC	AGT	ATT	TTA	CCT	AGC	AAT	GCG	AGC	AAC	AAG	ACG	GTG	GAT	TTA	ACC	1740	
Ser	Ser	Ile	Leu	Pro	Ser	Asn	Ala	Ser	Asn	Lys	Thr	Val	Asp	Leu	Thr		
		545					550					555					
CTA	ACG	CAA	TAC	CTT	TAT	GGC	TTA	GGG	GGG	CAA	GAG	TTA	TTG	TAT	TTT	1788	
Leu	Thr	Gln	Tyr	Leu	Tyr	Gly	Leu	Gly	Gly	Gln	Glu	Leu	Leu	Tyr	Phe		
	560					565					570						
AAA	ATA	TCG	CAA	CTC	ATC	AAT	CTT	GAC	GAT	AAA	GTT	TCG	CCC	TTT	AGA	1836	
Lys	Ile	Ser	Gln	Leu	Ile	Asn	Leu	Asp	Asp	Lys	Val	Ser	Pro	Phe	Arg		
575					580					585					590		
ATG	CCA	CTA	GAG	AGC	AAG	ATC	GGG	TTT	TCG	CCC	TTA	ACG	GGA	TTG	AAC	1884	
Met	Pro	Leu	Glu	Ser	Lys	Ile	Gly	Phe	Ser	Pro	Leu	Thr	Gly	Leu	Asn		
				595					600					605			
ATC	TTT	GGG	AAT	GTC	TTT	TAT	TCG	TTT	TAT	CAA	AAC	CGC	TTA	GAA	GAA	1932	
Ile	Phe	Gly	Asn	Val	Phe	Tyr	Ser	Phe	Tyr	Gln	Asn	Arg	Leu	Glu	Glu		
			610					615					620				
ATC	TCT	GTG	AAC	GCC	AAT	TAC	CAA	CGC	AAG	TTT	TTA	AGC	TTT	AAC	CTC	1980	
Ile	Ser	Val	Asn	Ala	Asn	Tyr	Gln	Arg	Lys	Phe	Leu	Ser	Phe	Asn	Leu		
			625				630					635					
TCT	TAT	TTT	TTA	AAA	AAC	AAT	TTT	AGC	AGT	GGG	ATT	AAT	AGC	ATT	GTA	2028	
Ser	Tyr	Phe	Leu	Lys	Asn	Asn	Phe	Ser	Ser	Gly	Ile	Asn	Ser	Ile	Val		
	640					645					650						
GAA	AAT	CTG	CGG	ATT	ATT	TAAAGGCGGG	TTTTAGCAAC	GACTTTGGCT	ATTTTTCC	2084							
Glu	Asn	Leu	Arg	Ile	Ile												
655					660												
ATGAGCGCGG	ATG															2097	

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 660 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

-131-

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```

Met Ile Tyr Trp Leu Tyr Leu Ala Val Phe Phe Leu Leu Ser Ala Leu
 1           5           10           15
Asp Ala Lys Glu Ile Ala Met Gln Arg Phe Asp Lys Gln Asn His Lys
 20           25           30
Ile Phe Glu Ile Leu Ala Asp Lys Val Ser Ala Lys Asp Asn Val Ile
 35           40           45
Thr Ala Ser Gly Asn Ala Ile Leu Leu Asn Tyr Asp Val Tyr Ile Leu
 50           55           60
Ala Asp Lys Val Arg Tyr Asp Thr Lys Thr Lys Glu Ala Leu Leu Glu
 65           70           75           80
Gly Asn Ile Lys Val Tyr Arg Gly Glu Gly Leu Leu Val Lys Thr Asp
 85           90           95
Tyr Val Lys Leu Ser Leu Asn Glu Lys Tyr Glu Ile Ile Phe Pro Phe
100           105           110
Tyr Val Gln Asp Ser Val Ser Gly Ile Trp Val Ser Ala Asp Ile Ala
115           120           125
Ser Gly Lys Asp Gln Lys Tyr Lys Val Lys Asn Met Ser Thr Ser Gly
130           135           140
Cys Ser Ile Asp Asn Pro Ile Trp His Val Asn Ala Thr Ser Gly Ser
145           150           155           160
Phe Asn Met Gln Lys Ser His Leu Ser Met Trp Asn Pro Lys Ile Tyr
165           170           175
Val Gly Asp Ile Pro Val Leu Tyr Leu Pro Tyr Ile Phe Met Ser Thr
180           185           190
Ser Asn Lys Arg Thr Thr Gly Phe Leu Tyr Pro Glu Phe Gly Thr Ser
195           200           205
Asn Leu Asp Gly Phe Ile Tyr Leu Gln Pro Phe Tyr Leu Ala Pro Lys
210           215           220
Asn Ser Trp Asp Met Thr Phe Thr Pro Gln Ile Arg Tyr Lys Arg Gly
225           230           235           240
Phe Gly Leu Asn Phe Glu Ala Arg Tyr Ile Asn Ser Lys Asn Asp Arg
245           250           255
Phe Leu Phe Asn Ala Arg Tyr Phe Arg Asn Tyr Thr Gln Tyr Val Lys
260           265           270
Arg Tyr Asp Leu Arg Asn Gln Asn Ile Tyr Gly Phe Glu Phe Leu Ser
275           280           285
Ser Ser Arg Asp Thr Leu Gln Lys Tyr Phe His Leu Lys Ser Asn Ile
290           295           300
Asp Asn Gly His Tyr Ile Asp Phe Leu Tyr Met Asn Asp Leu Asp Tyr
305           310           315           320
Val Arg Phe Glu Lys Val Asn Lys Arg Ile Thr Asp Ala Thr His Met
325           330           335
Ser Arg Ala Asn Tyr Tyr Leu Gln Thr Glu Asn Asn Tyr Tyr Gly Leu
340           345           350
Asn Ile Lys Tyr Phe Leu Asn Leu Asn Lys Ile Asn Asn Asn Arg Thr
355           360           365
Phe Gln Ser Val Pro Asn Leu Gln Tyr His Lys Tyr Leu Asn Ser Leu
370           375           380
Tyr Phe Arg Asn Leu Leu Tyr Ser Val Asp Tyr Gln Phe Arg Asn Thr
385           390           395           400

```

-132-

Ala Arg Glu Ile Gly Tyr Gly Tyr Val Gln Asn Ala Leu Asn Val Pro
405 410 415
Val Gly Leu Gln Phe Ser Leu Phe Lys Lys Tyr Leu Ser Leu Gly Leu
420 425 430
Trp Asn Asp Leu Gln Leu Ser Asn Val Ala Leu Met Gln Ser Lys Asn
435 440 445
Ser Phe Val Pro Thr Ile Pro Asn Glu Ser Arg Glu Phe Gly Asn Phe
450 455 460
Val Ser Ser Asn Phe Ser Met Tyr Val Asn Thr Asp Leu Ala Arg Glu
465 470 475 480
Tyr Asn Lys Leu Phe His Thr Ile Gln Leu Glu Ala Ile Phe Asn Ile
485 490 495
Pro Tyr Tyr Thr Phe Lys Asn Gly Leu Phe Ser Gln Asn Met Tyr Ala
500 505 510
Leu Ser Ala Gln Ala Leu Asn Ser Tyr Thr Ser Pro Leu Leu Arg Asp
515 520 525
Tyr Asp Tyr Gln Gly Arg Leu Tyr Asp Ser Val Trp Asn Pro Ser Ser
530 535 540
Ile Leu Pro Ser Asn Ala Ser Asn Lys Thr Val Asp Leu Thr Leu Thr
545 550 555 560
Gln Tyr Leu Tyr Gly Leu Gly Gly Gln Glu Leu Leu Tyr Phe Lys Ile
565 570 575
Ser Gln Leu Ile Asn Leu Asp Asp Lys Val Ser Pro Phe Arg Met Pro
580 585 590
Leu Glu Ser Lys Ile Gly Phe Ser Pro Leu Thr Gly Leu Asn Ile Phe
595 600 605
Gly Asn Val Phe Tyr Ser Phe Tyr Gln Asn Arg Leu Glu Glu Ile Ser
610 615 620
Val Asn Ala Asn Tyr Gln Arg Lys Phe Leu Ser Phe Asn Leu Ser Tyr
625 630 635 640
Phe Leu Lys Asn Asn Phe Ser Ser Gly Ile Asn Ser Ile Val Glu Asn
645 650 655
Leu Arg Ile Ile
660

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 961 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 168...764
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATGCCGATTA AATGCATGCT GATTAAATGA ATGAAAAGAG TCCAAACCAC CGCCTTTAAC	60
GCACCACGCT TGAAATTAAA ACTAAATTTT AGTGTATTCT TAGCAAATTT TAGATAAGAT	120

-133-

CAAGCGTGAT TTTTCTAAA TTTTAGGCAT TTAAGGAATC AGTGTTT ATG ACA AGC	176
Met Thr Ser	
1	
GCT CTG TTA GGC TTA CAA ATT GTT TTA GCG GTA TTG ATT GTG GTG GTG	224
Ala Leu Leu Gly Leu Gln Ile Val Leu Ala Val Leu Ile Val Val Val	
5 10 15	
GTT TTG TTG CAA AAA AGT TCT AGC ATC GGC TTA GGG GCT TAT AGC GGG	272
Val Leu Leu Gln Lys Ser Ser Ser Ile Gly Leu Gly Ala Tyr Ser Gly	
20 25 30 35	
AGT AAT GAG TCT TTA TTT GGC GCT AAA GGG CCT GCA AGC TTT ATG GCG	320
Ser Asn Glu Ser Leu Phe Gly Ala Lys Gly Pro Ala Ser Phe Met Ala	
40 45 50	
AAA TTA ACC ATG TTT TTA GGG CTG TTA TTT GTC ATC AAC ACC ATC GCT	368
Lys Leu Thr Met Phe Leu Gly Leu Leu Phe Val Ile Asn Thr Ile Ala	
55 60 65	
TTG GGC TAT TTT TAC AAC AAA GAA TAC GGC AAG AGC GTT TTA GAT GAG	416
Leu Gly Tyr Phe Tyr Asn Lys Glu Tyr Gly Lys Ser Val Leu Asp Glu	
70 75 80	
ACT AAA ACC AAC AAA GAA CTT TCG CCC CTA GTC CCT GCC ACC GGC ACG	464
Thr Lys Thr Asn Lys Glu Leu Ser Pro Leu Val Pro Ala Thr Gly Thr	
85 90 95	
CTT AAC CCT GCA CTT AAT CCC ACA TTA AAC CCA ACG CTC AAC CCT TTA	512
Leu Asn Pro Ala Leu Asn Pro Thr Leu Asn Pro Thr Leu Asn Pro Leu	
100 105 110 115	
GAG CAA GCC CCA ACT AAT CCT TTA ATG CCA CAA CAA ACG CCT AAC GAA	560
Glu Gln Ala Pro Thr Asn Pro Leu Met Pro Gln Gln Thr Pro Asn Glu	
120 125 130	
CTC CCT AAA GAG CCA GCC AAA ACG CCT TCT GTT GAA AGC CCC AAA CAG	608
Leu Pro Lys Glu Pro Ala Lys Thr Pro Ser Val Glu Ser Pro Lys Gln	
135 140 145	
AAT GAA AAG AAT GAA AAG AAT GAC GCC AAA GAG AAT GGT ATA AAG GGT	656
Asn Glu Lys Asn Glu Lys Asn Asp Ala Lys Glu Asn Gly Ile Lys Gly	
150 155 160	
GTT GAA AAA ACC AAA GAG AAC GCC AAA ACG CCC CCA ACC ACC CAC CAA	704
Val Glu Lys Thr Lys Glu Asn Ala Lys Thr Pro Pro Thr Thr His Gln	
165 170 175	
AAG CCT AAA ACG CAT GCA ACG CAA ACC AAC GCC CAT ACC AAC CAA AAA	752
Lys Pro Lys Thr His Ala Thr Gln Thr Asn Ala His Thr Asn Gln Lys	
180 185 190 195	
AAG GAT GAA AAA TAATGTTACA GGCCATTTAT AACGAAACCA AAGATCTGAT GCAAA	809
Lys Asp Glu Lys	

-134-

AAAGCATTCA	AGCTTTAAAC	AGGGATTTTT	CCACTCTAAG	GAGCGCGAAA	GTTTCAGTCA	869
ATATTTTAGA	TCACATCAAA	GTGGATTATT	ACGGCACGCC	CACGGCATT	AATCAAGTCG	929
GATCCGTGAT	GAGCTTGAT	GCGACCACCC	TT			961

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met	Thr	Ser	Ala	Leu	Leu	Gly	Leu	Gln	Ile	Val	Leu	Ala	Val	Leu	Ile	
1				5				10						15		
Val	Val	Val	Val	Leu	Leu	Gln	Lys	Ser	Ser	Ser	Ile	Gly	Leu	Gly	Ala	
			20				25						30			
Tyr	Ser	Gly	Ser	Asn	Glu	Ser	Leu	Phe	Gly	Ala	Lys	Gly	Pro	Ala	Ser	
		35				40					45					
Phe	Met	Ala	Lys	Leu	Thr	Met	Phe	Leu	Gly	Leu	Leu	Phe	Val	Ile	Asn	
	50					55				60						
Thr	Ile	Ala	Leu	Gly	Tyr	Phe	Tyr	Asn	Lys	Glu	Tyr	Gly	Lys	Ser	Val	
65					70				75					80		
Leu	Asp	Glu	Thr	Lys	Thr	Asn	Lys	Glu	Leu	Ser	Pro	Leu	Val	Pro	Ala	
			85					90						95		
Thr	Gly	Thr	Leu	Asn	Pro	Ala	Leu	Asn	Pro	Thr	Leu	Asn	Pro	Thr	Leu	
			100					105					110			
Asn	Pro	Leu	Glu	Gln	Ala	Pro	Thr	Asn	Pro	Leu	Met	Pro	Gln	Gln	Thr	
		115					120					125				
Pro	Asn	Glu	Leu	Pro	Lys	Glu	Pro	Ala	Lys	Thr	Pro	Ser	Val	Glu	Ser	
		130				135					140					
Pro	Lys	Gln	Asn	Glu	Lys	Asn	Glu	Lys	Asn	Asp	Ala	Lys	Glu	Asn	Gly	
145				150					155					160		
Ile	Lys	Gly	Val	Glu	Lys	Thr	Lys	Glu	Asn	Ala	Lys	Thr	Pro	Pro	Thr	
			165					170						175		
Thr	His	Gln	Lys	Pro	Lys	Thr	His	Ala	Thr	Gln	Thr	Asn	Ala	His	Thr	
			180				185						190			
Asn	Gln	Lys	Lys	Asp	Glu	Lys										
			195													

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1058 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence

-135-

(B) LOCATION: 325...879

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCTAGTCCCT	GCCACCGGCA	CGCTTAACCC	TGCACTTAAT	CCCACATTAA	ACCCAACGCT	60
CAACCCTTTA	GAGCAAGCCC	CAACTAATCC	TTTAATGCCA	CAACAAACGC	CTAACGAACT	120
CCCTAAAGAG	CCAGCCAAAA	CGCCTTCTGT	TGAAAGCCCC	AAACAGAATG	AAAAGAATGA	180
AAAGAATGAC	GCCAAAGAGA	ATGGTATAAA	GGGTGTTGAA	AAAACCAAAG	AGAACGCCAA	240
AACGCCCCCA	ACCACCCACC	AAAAGCCTAA	AACGCATGCA	ACGCAAACCA	ACGCCCATAC	300
CAACCAAAAA	AAGGATGAAA	AATA ATG TTA	CAG GCC ATT	TAT AAC GAA	ACC	351
		Met Leu Gln Ala	Ile Tyr Asn Glu Thr			
		1	5			
AAA GAT CTG	ATG CAA AAA	AGC ATT CAA	GCT TTA AAC	AGG GAT TTT	TCC	399
Lys Asp Leu	Met Gln Lys	Ser Ile Gln	Ala Leu Asn	Arg Asp Phe	Ser	
10	15	20	25			
ACT CTA AGG	AGC GCG AAA	GTT TCA GTC	AAT ATT TTA	GAT CAC ATC	AAA	447
Thr Leu Arg	Ser Ala Lys	Val Ser Val	Asn Ile Leu	Asp His Ile	Lys	
	30	35	40			
GTG GAT TAT	TAC GGC ACG	CCC ACG GCA	TTA AAT CAA	GTC GGA TCC	GTG	495
Val Asp Tyr	Tyr Gly Thr	Pro Thr Ala	Leu Asn Gln	Val Gly Ser	Val	
	45	50	55			
ATG AGC TTG	GAT GCG ACC	ACC CTT CAA	ATC AGC CCA	TGG GAA AAA	AAC	543
Met Ser Leu	Asp Ala Thr	Thr Leu Gln	Ile Ser Pro	Trp Glu Lys	Asn	
	60	65	70			
CTG CTC AAA	GAA ATT GAA	AGA TCC ATT	CAA GAA GCC	AAT ATT GGT	GTC	591
Leu Leu Lys	Glu Ile Glu	Arg Ser Ile	Gln Glu Ala	Asn Ile Gly	Val	
	75	80	85			
AAT CCT AAT	AAC GAC GGC	GAA ACG ATC	AAG CTT TTT	TTC CCG CCC	ATG	639
Asn Pro Asn	Asn Asp Gly	Glu Thr Ile	Lys Leu Phe	Phe Pro Pro	Met	
90	95	100	105			
ACA AGT GAG	CAA AGA AAA	CTC ATC GCA	AAA GAC GCC	AAA GCG ATG	GGT	687
Thr Ser Glu	Gln Arg Lys	Leu Ile Ala	Lys Asp Ala	Lys Ala Met	Gly	
	110	115	120			
GAA AAG GCT	AAA GTG GCT	GTG AGG AAT	ATC CGC CAA	GAT GCT AAC	AAC	735
Glu Lys Ala	Lys Val Ala	Val Arg Asn	Ile Arg Gln	Asp Ala Asn	Asn	
	125	130	135			
CAG GTG AAA	AAA TTA GAA	AAA GAC AAA	GAA ATC AGC	GAA GAT GAA	AGC	783
Gln Val Lys	Lys Leu Glu	Lys Asp Lys	Glu Ile Ser	Glu Asp Glu	Ser	
	140	145	150			
AAA AAA GCC	CAA GAG CAG	ATC CAA AAA	ATC ACC GAT	GAA GCC ATT	AAA	831
Lys Lys Ala	Gln Glu Gln	Ile Gln Lys	Ile Thr Asp	Glu Ala Ile	Lys	
	155	160	165			

-136-

```

AAA ATT GAT GAA AGC GTG AAA AAC AAA GAA GAC GCG ATC TTA AAG GTC T      880
Lys Ile Asp Glu Ser Val Lys Asn Lys Glu Asp Ala Ile Leu Lys Val
170                      175                      180                      185

AAACCATGGA TATTAAGGCA TGTTATCAAA ACGCTAAAGC GTTATTAGAG GGGCATTTC      940
TGCTCAGCAG TGGGTTTCAT TCCAATTATT ATTTGCAATC CGCTAAAGTT TTAGAAGATC      1000
CCAAACTAGC CGAACAATTA GCGCTAGAAT TAGCCAAACA AATCCAAGAA GCTCATTT      1058

```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 185 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```

Met Leu Gln Ala Ile Tyr Asn Glu Thr Lys Asp Leu Met Gln Lys Ser
 1             5             10             15
Ile Gln Ala Leu Asn Arg Asp Phe Ser Thr Leu Arg Ser Ala Lys Val
      20             25             30
Ser Val Asn Ile Leu Asp His Ile Lys Val Asp Tyr Tyr Gly Thr Pro
      35             40             45
Thr Ala Leu Asn Gln Val Gly Ser Val Met Ser Leu Asp Ala Thr Thr
      50             55             60
Leu Gln Ile Ser Pro Trp Glu Lys Asn Leu Leu Lys Glu Ile Glu Arg
      65             70             75             80
Ser Ile Gln Glu Ala Asn Ile Gly Val Asn Pro Asn Asn Asp Gly Glu
      85             90             95
Thr Ile Lys Leu Phe Phe Pro Pro Met Thr Ser Glu Gln Arg Lys Leu
      100            105            110
Ile Ala Lys Asp Ala Lys Ala Met Gly Glu Lys Ala Lys Val Ala Val
      115            120            125
Arg Asn Ile Arg Gln Asp Ala Asn Asn Gln Val Lys Lys Leu Glu Lys
      130            135            140
Asp Lys Glu Ile Ser Glu Asp Glu Ser Lys Lys Ala Gln Glu Gln Ile
      145            150            155            160
Gln Lys Ile Thr Asp Glu Ala Ile Lys Lys Ile Asp Glu Ser Val Lys
      165            170            175
Asn Lys Glu Asp Ala Ile Leu Lys Val
      180            185

```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

-137-

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 163...1389
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GAGTGGATGA	AAAAGACACT	TTCAATTTTG	CAAAAATTGG	CTATGAACAG	GGCAAGGGCG	60
AAGAATTAAA	AGAAGTAGAA	GAAAAGCATG	CGTTTAAGAA	AATCCCTTTT	GTCAAAGATT	120
TGCACAAAAT	CGCCCCCACT	ATCTTAAAAA	AGAGGCTATA	AA ATG GCT CAA AAT		174
				Met Ala Gln Asn		
				1		
TTC ACG AAA CTC AAC CCC CAG TTT GAA AAC ATC ATT TTT GAA CAT GAC						222
Phe Thr Lys Leu Asn Pro Gln Phe Glu Asn Ile Ile Phe Glu His Asp						
5 10 15 20						
GAC AAC CAA ATG ATT TTA AAC TTT GGC CCC CAA CAC CCC AGT AGT CAT						270
Asp Asn Gln Met Ile Leu Asn Phe Gly Pro Gln His Pro Ser Ser His						
25 30 35						
GGG CAA TTG CGC TTG ATT TTG GAA TTA GAG GGC GAA AAA ATC ATT AAG						318
Gly Gln Leu Arg Leu Ile Leu Glu Leu Glu Gly Glu Lys Ile Ile Lys						
40 45 50						
GCT ACC CCT GAA ATT GGC TAC TTG CAT AGA GGC TGT GAA AAG TTA GGC						366
Ala Thr Pro Glu Ile Gly Tyr Leu His Arg Gly Cys Glu Lys Leu Gly						
55 60 65						
GAA AAC ATG ACC TAT AAC GAA TAC ATG CCC ACT ACT GAT AGA TTG GAT						414
Glu Asn Met Thr Tyr Asn Glu Tyr Met Pro Thr Thr Asp Arg Leu Asp						
70 75 80						
TAC ACT TCT TCT ACC AGC AAT AAT TAC GCT TAC GCT TAT GCG GTA GAG						462
Tyr Thr Ser Ser Thr Ser Asn Asn Tyr Ala Tyr Ala Tyr Ala Val Glu						
85 90 95 100						
ACC TTA CTC AAT TTA GAA ATC CCA CGC CGA GCG CAG GTG ATC CGC ACG						510
Thr Leu Leu Asn Leu Glu Ile Pro Arg Arg Ala Gln Val Ile Arg Thr						
105 110 115						
ATT TTA CTA GAG CTT AAC CGC ATG ATC TCA CAC ATC TTT TTT ATC AGC						558
Ile Leu Leu Glu Leu Asn Arg Met Ile Ser His Ile Phe Phe Ile Ser						
120 125 130						
GTG CAT GCT TTA GAT GTG GGG GCG ATG AGC GTG TTT TTG TAT GCG TTT						606
Val His Ala Leu Asp Val Gly Ala Met Ser Val Phe Leu Tyr Ala Phe						
135 140 145						
AAA ACG AGG GAA TAC GGC TTG GAT TTG ATG GAG GAT TAT TGC GGG GCT						654
Lys Thr Arg Glu Tyr Gly Leu Asp Leu Met Glu Asp Tyr Cys Gly Ala						
150 155 160						
AGG CTC ACG CAT AAC GCT ATA AGG ATT GGG GGC GTG CCT TTA GAT TTA						702

-138-

Arg	Leu	Thr	His	Asn	Ala	Ile	Arg	Ile	Gly	Gly	Val	Pro	Leu	Asp	Leu		
165					170					175					180		
CCC	CCT	AAT	TGG	TTA	GAA	GGC	TTA	AAA	AAG	TTT	TTA	GGC	GAA	ATG	AGG	750	
Pro	Pro	Asn	Trp	Leu	Glu	Gly	Leu	Lys	Lys	Phe	Leu	Gly	Glu	Met	Arg		
				185					190					195			
GAA	TGC	AAA	AAA	CTC	ATT	CAA	GGC	TTA	TTG	GAT	AAG	AAT	CGC	ATT	TGG	798	
Glu	Cys	Lys	Lys	Leu	Ile	Gln	Gly	Leu	Leu	Asp	Lys	Asn	Arg	Ile	Trp		
			200					205					210				
CGG	ATG	CGC	TTG	GAA	AAT	GTG	GGC	GTT	GTA	ACG	CAA	AAA	ATG	GCG	CAA	846	
Arg	Met	Arg	Leu	Glu	Asn	Val	Gly	Val	Val	Thr	Gln	Lys	Met	Ala	Gln		
			215				220					225					
AGC	TGG	GGC	ATG	AGC	GGT	ATC	ATG	TTA	AGA	GGG	ACT	GGG	ATC	GCT	TAT	894	
Ser	Trp	Gly	Met	Ser	Gly	Ile	Met	Leu	Arg	Gly	Thr	Gly	Ile	Ala	Tyr		
	230					235					240						
GAC	ATC	AGA	AAA	GAA	GAG	CCT	TAT	GAG	CTT	TAT	AAA	GAG	CTT	GAT	TTT	942	
Asp	Ile	Arg	Lys	Glu	Glu	Pro	Tyr	Glu	Leu	Tyr	Lys	Glu	Leu	Asp	Phe		
	245				250					255					260		
GAT	GTG	CCG	GTG	GGC	AAT	TAT	GGC	GAT	AGT	TAT	GAT	AGG	TAT	TGT	TTG	990	
Asp	Val	Pro	Val	Gly	Asn	Tyr	Gly	Asp	Ser	Tyr	Asp	Arg	Tyr	Cys	Leu		
				265				270						275			
TAT	ATG	TTA	GAA	ATT	GAT	GAA	AGC	GTT	CGC	ATC	ATT	GAA	CAG	CTC	ATT	1038	
Tyr	Met	Leu	Glu	Ile	Asp	Glu	Ser	Val	Arg	Ile	Ile	Glu	Gln	Leu	Ile		
			280					285					290				
CCT	ATG	TAT	GCT	AAA	ACC	GAT	ACG	CCT	ATC	ATG	GCT	CAA	AAC	CCG	CAT	1086	
Pro	Met	Tyr	Ala	Lys	Thr	Asp	Thr	Pro	Ile	Met	Ala	Gln	Asn	Pro	His		
			295				300					305					
TAT	ATT	TCC	GCC	CCT	AAA	GAA	GAT	ATA	ATG	ACG	CAA	AAC	TAC	GCC	TTG	1134	
Tyr	Ile	Ser	Ala	Pro	Lys	Glu	Asp	Ile	Met	Thr	Gln	Asn	Tyr	Ala	Leu		
	310					315					320						
ATG	CAG	CAT	TTT	GTT	TTA	GTG	GCT	CAG	GGC	ATG	CGT	CCG	CCC	GTT	GGG	1182	
Met	Gln	His	Phe	Val	Leu	Val	Ala	Gln	Gly	Met	Arg	Pro	Pro	Val	Gly		
	325				330				335					340			
GAA	GTG	TAT	GCC	CCC	ACA	GAA	AGC	CCT	AAA	GGG	GAA	TTA	GGG	TTT	TTT	1230	
Glu	Val	Tyr	Ala	Pro	Thr	Glu	Ser	Pro	Lys	Gly	Glu	Leu	Gly	Phe	Phe		
				345					350					355			
ATC	CAT	TCA	GAG	GGC	GAG	CCT	TAC	CCT	CAC	AGG	CTA	AAA	ATC	AGA	GCC	1278	
Ile	His	Ser	Glu	Gly	Glu	Pro	Tyr	Pro	His	Arg	Leu	Lys	Ile	Arg	Ala		
			360					365					370				
CCT	AGC	TTT	TAT	CAC	ATT	GGG	GCT	TTG	AGC	GAC	ATT	TTA	GTG	GGG	CAA	1326	
Pro	Ser	Phe	Tyr	His	Ile	Gly	Ala	Leu	Ser	Asp	Ile	Leu	Val	Gly	Gln		
		375					380					385					

-139-

TAT TTA GCG GAT GCA GTA ACC GTG ATT GGC TCA ACC AAT GCG GTG TTT 1374
 Tyr Leu Ala Asp Ala Val Thr Val Ile Gly Ser Thr Asn Ala Val Phe
 390 395 400

GGC GAG GTG GAT AGA TGAAACGCTT TGATTACGC CCCTTAAAG CGGGTATTTT T 1430
 Gly Glu Val Asp Arg
 405

GAACGCTTAG AAGAATTGAT TGAAAAAGAA ATGCAACCTA ATGAAGTCGC TATTTTCATG 1490
 TTTGAAGTGG GGGATTTTTC TAATATCCCT AAGAGCGCTG AATTTATCCA ATCTAAAGGG 1550
 CATGAGCTCC TCAATTCCTT GCGTTTCAAT CAAGCGGATT GGACGATTGT CGTGAGAAAA 1610
 AAGGCTTGAT TTTGAGCGGC TTTAACCCT TAAATTCTCC CTTAGTCGCA AGCTCTTCT 1669

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 409 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Met Ala Gln Asn Phe Thr Lys Leu Asn Pro Gln Phe Glu Asn Ile Ile
 1 5 10 15
 Phe Glu His Asp Asp Asn Gln Met Ile Leu Asn Phe Gly Pro Gln His
 20 25 30
 Pro Ser Ser His Gly Gln Leu Arg Leu Ile Leu Glu Leu Glu Gly Glu
 35 40 45
 Lys Ile Ile Lys Ala Thr Pro Glu Ile Gly Tyr Leu His Arg Gly Cys
 50 55 60
 Glu Lys Leu Gly Glu Asn Met Thr Tyr Asn Glu Tyr Met Pro Thr Thr
 65 70 75 80
 Asp Arg Leu Asp Tyr Thr Ser Ser Thr Ser Asn Asn Tyr Ala Tyr Ala
 85 90 95
 Tyr Ala Val Glu Thr Leu Leu Asn Leu Glu Ile Pro Arg Arg Ala Gln
 100 105 110
 Val Ile Arg Thr Ile Leu Leu Glu Leu Asn Arg Met Ile Ser His Ile
 115 120 125
 Phe Phe Ile Ser Val His Ala Leu Asp Val Gly Ala Met Ser Val Phe
 130 135 140
 Leu Tyr Ala Phe Lys Thr Arg Glu Tyr Gly Leu Asp Leu Met Glu Asp
 145 150 155 160
 Tyr Cys Gly Ala Arg Leu Thr His Asn Ala Ile Arg Ile Gly Gly Val
 165 170 175
 Pro Leu Asp Leu Pro Pro Asn Trp Leu Glu Gly Leu Lys Lys Phe Leu
 180 185 190
 Gly Glu Met Arg Glu Cys Lys Lys Leu Ile Gln Gly Leu Leu Asp Lys
 195 200 205
 Asn Arg Ile Trp Arg Met Arg Leu Glu Asn Val Gly Val Val Thr Gln
 210 215 220
 Lys Met Ala Gln Ser Trp Gly Met Ser Gly Ile Met Leu Arg Gly Thr

-140-

```

225          230          235          240
Gly Ile Ala Tyr Asp Ile Arg Lys Glu Glu Pro Tyr Glu Leu Tyr Lys
          245          250          255
Glu Leu Asp Phe Asp Val Pro Val Gly Asn Tyr Gly Asp Ser Tyr Asp
          260          265          270
Arg Tyr Cys Leu Tyr Met Leu Glu Ile Asp Glu Ser Val Arg Ile Ile
          275          280          285
Glu Gln Leu Ile Pro Met Tyr Ala Lys Thr Asp Thr Pro Ile Met Ala
          290          295          300
Gln Asn Pro His Tyr Ile Ser Ala Pro Lys Glu Asp Ile Met Thr Gln
305          310          315          320
Asn Tyr Ala Leu Met Gln His Phe Val Leu Val Ala Gln Gly Met Arg
          325          330          335
Pro Pro Val Gly Glu Val Tyr Ala Pro Thr Glu Ser Pro Lys Gly Glu
          340          345          350
Leu Gly Phe Phe Ile His Ser Glu Gly Glu Pro Tyr Pro His Arg Leu
          355          360          365
Lys Ile Arg Ala Pro Ser Phe Tyr His Ile Gly Ala Leu Ser Asp Ile
          370          375          380
Leu Val Gly Gln Tyr Leu Ala Asp Ala Val Thr Val Ile Gly Ser Thr
385          390          395          400
Asn Ala Val Phe Gly Glu Val Asp Arg
          405

```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 869 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 358...732
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

TAACTTGTGG TTAACCTACCG CCAGACTCCT TTTGAGTTTG GCAAACGCGC CAATGAGTTC      60
TTTAGGCATT TTTTCAGTGC CGATCTTAAA GTTTTCAAGA CTGCGTTGCG TTTGAGCCCC      120
CCAATATTGG CTATCATTTA CTTTGATTTT GCCCATCGTG TCATGTTCAA TTCTAAATTG      180
CATGCTAATC CTTTGAAATT TGATTTTAAA ACCTTAAAAA AATAGCATAA ACTCTTATAC      240
CTTCTACTTA AAAACCCTAA TTTTAAAC ACCATTTCCA CAATTTTAC ACAAAGAGG      300
GTTATTATCC GTTCGCAACA AGAATTTTCT TGTTATCTTA ATGTAAAGGT CAAAACG ATG      360
                                         Met
                                         1

```

```

AAA AAG TTA GCC GCT TTA TTT TTA GTA AGC GTG TTG GGG GTT ATG GGT      408
Lys Lys Leu Ala Ala Leu Phe Leu Val Ser Val Leu Gly Val Met Gly
          5          10          15

```

-141-

TTA AAC GCA TGG GAG CAA ACC CTA AAA GCT AAT GAC TTG GAA GTG AAA	456
Leu Asn Ala Trp Glu Gln Thr Leu Lys Ala Asn Asp Leu Glu Val Lys	
20 25 30	
ATC AAA TCC GTG GGT AAC CCC ATT AAA GGC GAT AAC ACT TTC ATT CTC	504
Ile Lys Ser Val Gly Asn Pro Ile Lys Gly Asp Asn Thr Phe Ile Leu	
35 40 45	
AGC CCC ACT TTA AAA GGT AAG GCT TTA GAA AAA GCT ATC GTT AGG GTG	552
Ser Pro Thr Leu Lys Gly Lys Ala Leu Glu Lys Ala Ile Val Arg Val	
50 55 60 65	
CAG TTT ATG ATG CCT GAA ATG CCC GGC ATG CCA GCG ATG AAA GAA ATG	600
Gln Phe Met Met Pro Glu Met Pro Gly Met Pro Ala Met Lys Glu Met	
70 75 80	
GCG CAA GTG AGT GAA AAA AAC GGC CTT TAT GAA GCT AAA ACC AAT CTT	648
Ala Gln Val Ser Glu Lys Asn Gly Leu Tyr Glu Ala Lys Thr Asn Leu	
85 90 95	
TCT ATG AAC GGG ACA TGG CAG GTT AGG GTG GAT ATT AAA TCT AAA GAG	696
Ser Met Asn Gly Thr Trp Gln Val Arg Val Asp Ile Lys Ser Lys Glu	
100 105 110	
GGT CAG GTT TAT CGC GCT AAA ACA AGC CTG GAT TTA TAAGAGCATG CTATCT	748
Gly Gln Val Tyr Arg Ala Lys Thr Ser Leu Asp Leu	
115 120 125	
TTTATAAGCG CGTTTGATAA AAGGGGCGTT TCAATACGCC TTTTAACAGC CTTGTTACTG	808
CTTTTTAGTT TGGGTTTGGC TAAAGATTTA GAGATCCAAT CTTTTGTGGC TAAATACCTT	868
T	869

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Lys Lys Leu Ala Ala Leu Phe Leu Val Ser Val Leu Gly Val Met	
1 5 10 15	
Gly Leu Asn Ala Trp Glu Gln Thr Leu Lys Ala Asn Asp Leu Glu Val	
20 25 30	
Lys Ile Lys Ser Val Gly Asn Pro Ile Lys Gly Asp Asn Thr Phe Ile	
35 40 45	
Leu Ser Pro Thr Leu Lys Gly Lys Ala Leu Glu Lys Ala Ile Val Arg	
50 55 60	
Val Gln Phe Met Met Pro Glu Met Pro Gly Met Pro Ala Met Lys Glu	
65 70 75 80	

-142-

Met Ala Gln Val Ser Glu Lys Asn Gly Leu Tyr Glu Ala Lys Thr Asn
 85 90 95
 Leu Ser Met Asn Gly Thr Trp Gln Val Arg Val Asp Ile Lys Ser Lys
 100 105 110
 Glu Gly Gln Val Tyr Arg Ala Lys Thr Ser Leu Asp Leu
 115 120 125

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1217 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 73...1152
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCCATGCGTT TTGATGCGAT TTTAAAAAAT CTTTGGGTAT TTTAGCATGC CAATGGTTAA	60
AAAAAGGTGG TT ATG AAT GGT TTT TGC GCT AGA CTA CGA GCC ATA ACT CAT	111
Met Asn Gly Phe Cys Ala Arg Leu Arg Ala Ile Thr His	
1 5 10	
AAT GAA AGA TTA AAA ATG AAA ATA GCG GTA TTA CTC AGT GGG GGG GTG	159
Asn Glu Arg Leu Lys Met Lys Ile Ala Val Leu Leu Ser Gly Gly Val	
15 20 25	
GAT AGC TCT TAT AGC GCT TAT AGC TTA AAA GAG CAA GGG CAT GAA TTA	207
Asp Ser Ser Tyr Ser Ala Tyr Ser Leu Lys Glu Gln Gly His Glu Leu	
30 35 40 45	
GTG GGG ATT TAT TTA AAA CTC CAT GCG AGT GAA AAA AAG CAT GAT TTA	255
Val Gly Ile Tyr Leu Lys Leu His Ala Ser Glu Lys Lys His Asp Leu	
50 55 60	
TAC ATC AAA AAC GCT CAA AAA GCA TGC GAG TTT TTA GGC ATT CCT TTA	303
Tyr Ile Lys Asn Ala Gln Lys Ala Cys Glu Phe Leu Gly Ile Pro Leu	
65 70 75	
GAG GTG TTG GAT TTT CAA AAG GAT TTT AAA AGC GCG GTT TAT GAT GAA	351
Glu Val Leu Asp Phe Gln Lys Asp Phe Lys Ser Ala Val Tyr Asp Glu	
80 85 90	
TTT ATC AAC GCC TAT GAA GAA GGG CAA ACC CCA AAC CCT TGT GCG TTG	399
Phe Ile Asn Ala Tyr Glu Glu Gly Gln Thr Pro Asn Pro Cys Ala Leu	
95 100 105	
TGC AAC CCT TTA ATG AAG TTT GGG CTA GCT TTG GAT CAC GCT TTA AAA	447

-143-

Cys	Asn	Pro	Leu	Met	Lys	Phe	Gly	Leu	Ala	Leu	Asp	His	Ala	Leu	Lys	
110					115					120					125	
TTA	GGG	TGT	GAA	AAG	ATC	GCT	ACC	GGG	CAT	TAT	GCG	AGA	GTC	AAA	GAA	495
Leu	Gly	Cys	Glu	Lys	Ile	Ala	Thr	Gly	His	Tyr	Ala	Arg	Val	Lys	Glu	
				130					135					140		
ATT	GAC	AAA	ATA	AGT	TAT	ATT	CAA	GAG	GCT	TTG	GAT	AAA	ACT	AAA	GAT	543
Ile	Asp	Lys	Ile	Ser	Tyr	Ile	Gln	Glu	Ala	Leu	Asp	Lys	Thr	Lys	Asp	
			145					150					155			
CAG	AGC	TAT	TTT	TTA	TAC	GCT	TTA	GAG	CAT	GAA	GTG	ATC	GCT	AAA	TTG	591
Gln	Ser	Tyr	Phe	Leu	Tyr	Ala	Leu	Glu	His	Glu	Val	Ile	Ala	Lys	Leu	
		160					165					170				
GTG	TTC	CCT	TTA	GGG	GAT	TTG	CTA	AAA	AAG	GAT	ATT	AAG	CCT	TTA	GCC	639
Val	Phe	Pro	Leu	Gly	Asp	Leu	Leu	Lys	Lys	Asp	Ile	Lys	Pro	Leu	Ala	
	175					180					185					
TTG	AAT	GCG	ATG	CCT	TTT	TTA	GGC	ACT	TTA	GAG	ACT	TAT	AAG	GAA	TCT	687
Leu	Asn	Ala	Met	Pro	Phe	Leu	Gly	Thr	Leu	Glu	Thr	Tyr	Lys	Glu	Ser	
190					195					200					205	
CAA	GAA	ATC	TGC	TTT	GTG	GAA	AAA	AGC	TAC	ATT	GAC	ACT	TTA	AAA	AAG	735
Gln	Glu	Ile	Cys	Phe	Val	Glu	Lys	Ser	Tyr	Ile	Asp	Thr	Leu	Lys	Lys	
				210					215					220		
CAT	GTT	GAA	GTG	GAA	AAA	GAG	GGC	GTG	GTG	AAA	AAC	CTA	CAA	GGC	GAA	783
His	Val	Glu	Val	Glu	Lys	Glu	Gly	Val	Val	Lys	Asn	Leu	Gln	Gly	Glu	
			225					230					235			
GTC	ATT	GGC	ACG	CAT	AAA	GGC	TAT	ATG	CAA	TAC	ACG	ATT	GGC	AAA	CGC	831
Val	Ile	Gly	Thr	His	Lys	Gly	Tyr	Met	Gln	Tyr	Thr	Ile	Gly	Lys	Arg	
	240					245						250				
AAA	GGC	TTT	AGT	ATT	AAA	GGC	GCG	TTA	GAG	CCG	CAT	TTT	GTG	GTG	GGG	879
Lys	Gly	Phe	Ser	Ile	Lys	Gly	Ala	Leu	Glu	Pro	His	Phe	Val	Val	Gly	
	255					260				265						
ATT	GAC	GCT	AAA	AAG	AAC	GAG	CTA	GTC	GTG	GGC	AAA	AAA	GAA	GAT	CTC	927
Ile	Asp	Ala	Lys	Lys	Asn	Glu	Leu	Val	Val	Gly	Lys	Lys	Glu	Asp	Leu	
270					275					280					285	
GCC	ACG	CAT	TCG	CTT	AAG	GCT	AAA	AAC	AAA	TCT	TTA	ATG	AAA	GAT	TTT	975
Ala	Thr	His	Ser	Leu	Lys	Ala	Lys	Asn	Lys	Ser	Leu	Met	Lys	Asp	Phe	
				290					295					300		
AAA	GAT	GGC	GAA	TAT	TTT	ATC	AAG	GCT	CGT	TAC	AGG	AGC	GTG	CCT	GCT	1023
Lys	Asp	Gly	Glu	Tyr	Phe	Ile	Lys	Ala	Arg	Tyr	Arg	Ser	Val	Pro	Ala	
			305					310					315			
AAA	GCG	CAT	GTG	AGT	TTG	AAA	GAT	GAG	GTG	ATT	GAA	GTG	GGG	TTT	AAA	1071
Lys	Ala	His	Val	Ser	Leu	Lys	Asp	Glu	Val	Ile	Glu	Val	Gly	Phe	Lys	
			320				325						330			

-144-

GAG CCT TTT TAT GGC GTG GCT AAA GGG CAA GCT TTG GTC GTT TAT AAA 1119
 Glu Pro Phe Tyr Gly Val Ala Lys Gly Gln Ala Leu Val Val Tyr Lys
 335 340 345

GAT GAC ATC TTG CTT GGT GGG GGC GTG ATT GTT TAAAACTAA AGAACTAAGA 1172
 Asp Asp Ile Leu Leu Gly Gly Gly Val Ile Val
 350 355 360

GATACGCCTT TTGGCAGTCT CTTAATGTTT TATTGAATAG GCGTT 1217

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met	Asn	Gly	Phe	Cys	Ala	Arg	Leu	Arg	Ala	Ile	Thr	His	Asn	Glu	Arg	1	5	10	15
Leu	Lys	Met	Lys	Ile	Ala	Val	Leu	Leu	Ser	Gly	Gly	Val	Asp	Ser	Ser	20	25	30	
Tyr	Ser	Ala	Tyr	Ser	Leu	Lys	Glu	Gln	Gly	His	Glu	Leu	Val	Gly	Ile	35	40	45	
Tyr	Leu	Lys	Leu	His	Ala	Ser	Glu	Lys	Lys	His	Asp	Leu	Tyr	Ile	Lys	50	55	60	
Asn	Ala	Gln	Lys	Ala	Cys	Glu	Phe	Leu	Gly	Ile	Pro	Leu	Glu	Val	Leu	65	70	75	80
Asp	Phe	Gln	Lys	Asp	Phe	Lys	Ser	Ala	Val	Tyr	Asp	Glu	Phe	Ile	Asn	85	90	95	
Ala	Tyr	Glu	Glu	Gly	Gln	Thr	Pro	Asn	Pro	Cys	Ala	Leu	Cys	Asn	Pro	100	105	110	
Leu	Met	Lys	Phe	Gly	Leu	Ala	Leu	Asp	His	Ala	Leu	Lys	Leu	Gly	Cys	115	120	125	
Glu	Lys	Ile	Ala	Thr	Gly	His	Tyr	Ala	Arg	Val	Lys	Glu	Ile	Asp	Lys	130	135	140	
Ile	Ser	Tyr	Ile	Gln	Glu	Ala	Leu	Asp	Lys	Thr	Lys	Asp	Gln	Ser	Tyr	145	150	155	160
Phe	Leu	Tyr	Ala	Leu	Glu	His	Glu	Val	Ile	Ala	Lys	Leu	Val	Phe	Pro	165	170	175	
Leu	Gly	Asp	Leu	Leu	Lys	Lys	Asp	Ile	Lys	Pro	Leu	Ala	Leu	Asn	Ala	180	185	190	
Met	Pro	Phe	Leu	Gly	Thr	Leu	Glu	Thr	Tyr	Lys	Glu	Ser	Gln	Glu	Ile	195	200	205	
Cys	Phe	Val	Glu	Lys	Ser	Tyr	Ile	Asp	Thr	Leu	Lys	Lys	His	Val	Glu	210	215	220	
Val	Glu	Lys	Glu	Gly	Val	Val	Lys	Asn	Leu	Gln	Gly	Glu	Val	Ile	Gly	225	230	235	240
Thr	His	Lys	Gly	Tyr	Met	Gln	Tyr	Thr	Ile	Gly	Lys	Arg	Lys	Gly	Phe	245	250	255	

-145-

```

Ser Ile Lys Gly Ala Leu Glu Pro His Phe Val Val Gly Ile Asp Ala
      260                      265                      270
Lys Lys Asn Glu Leu Val Val Gly Lys Lys Glu Asp Leu Ala Thr His
      275                      280                      285
Ser Leu Lys Ala Lys Asn Lys Ser Leu Met Lys Asp Phe Lys Asp Gly
      290                      295                      300
Glu Tyr Phe Ile Lys Ala Arg Tyr Arg Ser Val Pro Ala Lys Ala His
      305                      310                      315                      320
Val Ser Leu Lys Asp Glu Val Ile Glu Val Gly Phe Lys Glu Pro Phe
      325                      330                      335
Tyr Gly Val Ala Lys Gly Gln Ala Leu Val Val Tyr Lys Asp Asp Ile
      340                      345                      350
Leu Leu Gly Gly Gly Val Ile Val
      355                      360

```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 975 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 191...793
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```

ACATTACACA TATCTGTCGC TAAAACGCGC CGCTTCACTA AACCCACTGA TTGTAAAAAT      60
TTGTCTATTC GCATGCGTTT ATTTTACCCT ATTCTTTAAG TTTTATATCCA TAACTTATAA      120
GGGTTTTAGT TTTAGCATGT TAGCATTCAG CCACCACTCT TTTTAAGGAA TTTGTTTGAA      180
GTTTCAAATT ATG AGT TTG TTA GCC ACT CTT TTA TTA GCC TCT TGC TTG      229
          Met Ser Leu Leu Ala Thr Leu Leu Leu Ala Ser Cys Leu
              1              5              10

CCC CCC AAA GGC CAT CAT TCT GGT TTG GTG AAT CTT TAT ATC GCT CAT      277
Pro Pro Lys Gly His His Ser Gly Leu Val Asn Leu Tyr Ile Ala His
      15              20              25

CAA GGC CAA AGC GTG CGC ACT TAT TGG CGC AAA GTG GAT AGA GGA GTT      325
Gln Gly Gln Ser Val Arg Thr Tyr Trp Arg Lys Val Asp Arg Gly Val
      30              35              40              45

ATC GCT AAA CAC AAT GAA GCG CTT AAA AAA GAT CCT AAA GCA AAG CTC      373
Ile Ala Lys His Asn Glu Ala Leu Lys Lys Asp Pro Lys Ala Lys Leu
              50              55              60

AAA GAC CCC AGG GGG CCT TTA TTC ATG CTA GGG AGT GAG CGC TTC ATG      421
Lys Asp Pro Arg Gly Pro Leu Phe Met Leu Gly Ser Glu Arg Phe Met
      65              70              75

```

-146-

CTT TTA TGG AAA AAC CGC TAC GCT TTA GCC AAG CCC CAA TCG TTC AGG	469
Leu Leu Trp Lys Asn Arg Tyr Ala Leu Ala Lys Pro Gln Ser Phe Arg	
80 85 90	
CTA GAG CCT GGT TTT TAT TAC TTG GAT TCT TTT AGC GTG GAA ACT CAA	517
Leu Glu Pro Gly Phe Tyr Tyr Leu Asp Ser Phe Ser Val Glu Thr Gln	
95 100 105	
AAA GGC GTC TTG CAG AGC GCT CCT GGC TAT TCA TAT ACT AAA AAT GGC	565
Lys Gly Val Leu Gln Ser Ala Pro Gly Tyr Ser Tyr Thr Lys Asn Gly	
110 115 120 125	
TAT GAT TTC AAA AAC AAC CGC CCC TTT TTC CTG GCC TTT GAA GTC AAA	613
Tyr Asp Phe Lys Asn Asn Arg Pro Phe Phe Leu Ala Phe Glu Val Lys	
130 135 140	
CCT GAT GGC AAA ACC ATT CTT CCT AGC GTG GAA TTA AGC CTG ATT AAA	661
Pro Asp Gly Lys Thr Ile Leu Pro Ser Val Glu Leu Ser Leu Ile Lys	
145 150 155	
ACC CCT AGA GGC TTT TTA GGG GTG TTC TTG TTT GAT AAT AAT GAA AAG	709
Thr Pro Arg Gly Phe Leu Gly Val Phe Leu Phe Asp Asn Asn Glu Lys	
160 165 170	
GGG ACT AAC GCC AAG TGG ATT GAG GGG AGT TTG AAT TTA AAG CTT AAA	757
Gly Thr Asn Ala Lys Trp Ile Glu Gly Ser Leu Asn Leu Lys Leu Lys	
175 180 185	
AAC GCT TCC TTT AAA GAT GCG TGG GGG TTG GAA CAA TAAAGCATGA AGTGAT	809
Asn Ala Ser Phe Lys Asp Ala Trp Gly Leu Glu Gln	
190 195 200	
CGCTTGCTTT TCGTAAGCTC TTTATGATTA GATTGTAAAA AAATGCCTTG AGTATTTTTT	869
AGATTTTATT ACCCCTATTC AATTGGAACA AAGCCATTAA ATTTTTAAAA ACTTTTAAAA	929
ACGATAAACA TAATCCGCGC TCCAAGTAAC ATAGCTTTCA AAAATG	975

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 201 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Ser Leu Leu Ala Thr Leu Leu Leu Ala Ser Cys Leu Pro Pro Lys
1 5 10 15
Gly His His Ser Gly Leu Val Asn Leu Tyr Ile Ala His Gln Gly Gln
20 25 30
Ser Val Arg Thr Tyr Trp Arg Lys Val Asp Arg Gly Val Ile Ala Lys
35 40 45
His Asn Glu Ala Leu Lys Lys Asp Pro Lys Ala Lys Leu Lys Asp Pro

-147-

50		55		60
Arg Gly Pro Leu Phe Met Leu Gly Ser Glu Arg Phe Met Leu Leu Trp				
65		70		75
Lys Asn Arg Tyr Ala Leu Ala Lys Pro Gln Ser Phe Arg Leu Glu Pro				80
	85		90	95
Gly Phe Tyr Tyr Leu Asp Ser Phe Ser Val Glu Thr Gln Lys Gly Val				
	100		105	110
Leu Gln Ser Ala Pro Gly Tyr Ser Tyr Thr Lys Asn Gly Tyr Asp Phe				
	115		120	125
Lys Asn Asn Arg Pro Phe Phe Leu Ala Phe Glu Val Lys Pro Asp Gly				
	130		135	140
Lys Thr Ile Leu Pro Ser Val Glu Leu Ser Leu Ile Lys Thr Pro Arg				
145		150		155
Gly Phe Leu Gly Val Phe Leu Phe Asp Asn Asn Glu Lys Gly Thr Asn				160
	165		170	175
Ala Lys Trp Ile Glu Gly Ser Leu Asn Leu Lys Leu Lys Asn Ala Ser				
	180		185	190
Phe Lys Asp Ala Trp Gly Leu Glu Gln				
	195		200	

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1116 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 90...1076
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TATAAATACA TCGTTTCATT AGCGAATTTA ATGGCGTTAA GCGATCATAT TGATTTATTT	60
TATGAATTTG TTTATTAAGG GAAAAAATC ATG TCA AAT AGC ATG TTG GAT AAA	113
Met Ser Asn Ser Met Leu Asp Lys	
1 5	
AAT AAA GCG ATT CTT ACA GGG GGT GGG GCT TTA TTA TTA GGG CTA ATC	161
Asn Lys Ala Ile Leu Thr Gly Gly Gly Ala Leu Leu Leu Gly Leu Ile	
10 15 20	
GTG CTT TTT TAT TTA GCT TAT CGC CCT AAG GCT GAA GTG TTG CAA GGG	209
Val Leu Phe Tyr Leu Ala Tyr Arg Pro Lys Ala Glu Val Leu Gln Gly	
25 30 35 40	
TTT TTG GAA GCC AGA GAA TAC AGC GTG AGC TCC AAA GTC CCT GGC CGC	257
Phe Leu Glu Ala Arg Glu Tyr Ser Val Ser Ser Lys Val Pro Gly Arg	
45 50 55	

-148-

ATT GAA AAG GTG TTT GTT AAA AAA GGC GAT CAC ATT AAA AAG GGC GAT	305
Ile Glu Lys Val Phe Val Lys Lys Gly Asp His Ile Lys Lys Gly Asp	
60 65 70	
TTG GTT TTT AGC ATT TCT AGC CCT GAA TTA GAA GCC AAA CTC GCT CAA	353
Leu Val Phe Ser Ile Ser Ser Pro Glu Leu Glu Ala Lys Leu Ala Gln	
75 80 85	
GCT GAA GCC GGG CAT AAA GCC GCT AAA GCG CTT AGC GAT GAA GTC AAA	401
Ala Glu Ala Gly His Lys Ala Ala Lys Ala Leu Ser Asp Glu Val Lys	
90 95 100	
AGA GGC TCA AGA GAC GAA ACG ATT AAT TCT GCG AGA GAC GTT TGG CAA	449
Arg Gly Ser Arg Asp Glu Thr Ile Asn Ser Ala Arg Asp Val Trp Gln	
105 110 115 120	
GCA GCC AAA TCC CAA GCC ACT TTA GCC AAA GAG ACT TAT AAG CGC GTT	497
Ala Ala Lys Ser Gln Ala Thr Leu Ala Lys Glu Thr Tyr Lys Arg Val	
125 130 135	
CAA GAT TTG TAT GAT AAT GGC GTG GCG AGC TTG CAA AAG CGC GAT GAA	545
Gln Asp Leu Tyr Asp Asn Gly Val Ala Ser Leu Gln Lys Arg Asp Glu	
140 145 150	
GCC TAT GCG GCT TAT GAA AGC ACT AAA TAC AAC GAG AGC GCG GCT TAC	593
Ala Tyr Ala Ala Tyr Glu Ser Thr Lys Tyr Asn Glu Ser Ala Ala Tyr	
155 160 165	
CAA AAG TAT AAA ATG GCT TTA GGG GGG GCG AGC TCT GAA AGT AAG ATT	641
Gln Lys Tyr Lys Met Ala Leu Gly Gly Ala Ser Ser Glu Ser Lys Ile	
170 175 180	
GCC GCT AAG GCT AAA GAG AGC GCG GCT TTA GGG CAA GTG AAT GAA GTG	689
Ala Ala Lys Ala Lys Glu Ser Ala Ala Leu Gly Gln Val Asn Glu Val	
185 190 195 200	
GAG TCT TAT TTA AAA GAC GTC AAA GCG ACA GCC CCA ATT GAT GGG GAA	737
Glu Ser Tyr Leu Lys Asp Val Lys Ala Thr Ala Pro Ile Asp Gly Glu	
205 210 215	
GTG AGT AAC GTG CTT TTA AGC GGT GGC GAG CTT AGC CCT AAG GGT TTT	785
Val Ser Asn Val Leu Leu Ser Gly Gly Glu Leu Ser Pro Lys Gly Phe	
220 225 230	
CCT GTG GTT TTA ATG ATA GAT TTA AAG GAT AGT TGG TTA AAA ATC AGC	833
Pro Val Val Leu Met Ile Asp Leu Lys Asp Ser Trp Leu Lys Ile Ser	
235 240 245	
GTG CCT GAA AAG TAT TTG AAC GAG TTT AAA GTG GGT AAG GAA TTT GAA	881
Val Pro Glu Lys Tyr Leu Asn Glu Phe Lys Val Gly Lys Glu Phe Glu	
250 255 260	
GGC TAT ATC CCG GCG TTG AAA AAA AGC ACG AAA TTC AGG GTC AAA TAT	929
Gly Tyr Ile Pro Ala Leu Lys Lys Ser Thr Lys Phe Arg Val Lys Tyr	
265 270 275 280	

-149-

TTG AGC GTG ATG GGG GAT TTT GCG ACT TGG AAA GCG ACG AAT AAT TCC	977
Leu Ser Val Met Gly Asp Phe Ala Thr Trp Lys Ala Thr Asn Asn Ser	
285 290 295	
AAC ACT TAC GAC ATG AAA AGC TAT GAA GTG GAA GCC ATA CCC TTA GAA	1025
Asn Thr Tyr Asp Met Lys Ser Tyr Glu Val Glu Ala Ile Pro Leu Glu	
300 305 310	
GAG TTG GAA AAT TTT AGG GTA GGG ATG AGC GTG TTA GTT ACC ATT AAA	1073
Glu Leu Glu Asn Phe Arg Val Gly Met Ser Val Leu Val Thr Ile Lys	
315 320 325	
CCT TAAAAAGGAT TGTTTTGTTC AGATTGATAA GCGCATGGGT	1116
Pro	

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Met Ser Asn Ser Met Leu Asp Lys Asn Lys Ala Ile Leu Thr Gly Gly	
1 5 10 15	
Gly Ala Leu Leu Leu Gly Leu Ile Val Leu Phe Tyr Leu Ala Tyr Arg	
20 25 30	
Pro Lys Ala Glu Val Leu Gln Gly Phe Leu Glu Ala Arg Glu Tyr Ser	
35 40 45	
Val Ser Ser Lys Val Pro Gly Arg Ile Glu Lys Val Phe Val Lys Lys	
50 55 60	
Gly Asp His Ile Lys Lys Gly Asp Leu Val Phe Ser Ile Ser Ser Pro	
65 70 75 80	
Glu Leu Glu Ala Lys Leu Ala Gln Ala Glu Ala Gly His Lys Ala Ala	
85 90 95	
Lys Ala Leu Ser Asp Glu Val Lys Arg Gly Ser Arg Asp Glu Thr Ile	
100 105 110	
Asn Ser Ala Arg Asp Val Trp Gln Ala Ala Lys Ser Gln Ala Thr Leu	
115 120 125	
Ala Lys Glu Thr Tyr Lys Arg Val Gln Asp Leu Tyr Asp Asn Gly Val	
130 135 140	
Ala Ser Leu Gln Lys Arg Asp Glu Ala Tyr Ala Ala Tyr Glu Ser Thr	
145 150 155 160	
Lys Tyr Asn Glu Ser Ala Ala Tyr Gln Lys Tyr Lys Met Ala Leu Gly	
165 170 175	
Gly Ala Ser Ser Glu Ser Lys Ile Ala Ala Lys Ala Lys Glu Ser Ala	
180 185 190	
Ala Leu Gly Gln Val Asn Glu Val Glu Ser Tyr Leu Lys Asp Val Lys	

-150-

195	200	205
Ala Thr Ala Pro Ile Asp Gly Glu Val Ser Asn Val Leu Leu Ser Gly		
210	215	220
Gly Glu Leu Ser Pro Lys Gly Phe Pro Val Val Leu Met Ile Asp Leu		
225	230	235
Lys Asp Ser Trp Leu Lys Ile Ser Val Pro Glu Lys Tyr Leu Asn Glu		240
	245	250
Phe Lys Val Gly Lys Glu Phe Glu Gly Tyr Ile Pro Ala Leu Lys Lys		255
	260	265
Ser Thr Lys Phe Arg Val Lys Tyr Leu Ser Val Met Gly Asp Phe Ala		270
	275	280
Thr Trp Lys Ala Thr Asn Asn Ser Asn Thr Tyr Asp Met Lys Ser Tyr		285
	290	295
Glu Val Glu Ala Ile Pro Leu Glu Glu Leu Glu Asn Phe Arg Val Gly		300
305	310	315
Met Ser Val Leu Val Thr Ile Lys Pro		320
	325	

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1514 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 94...1467
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

AAATAAAATA GCGATCATTA TAACATGTTG CTTTTTAAGT GAAAGCGTTA AGTTGTTAGG	60
GTATAGTGGC TTAAAAATTT TAGGATATTG AGA ATG CTT GAA ACT TCT AGC CAT	114
Met Leu Glu Thr Ser Ser His	
1 5	
TTT TTA AAA TCG TTT CGC TTG AAG CGT TAT ATA GGG TTT TTA TTG ATT	162
Phe Leu Lys Ser Phe Arg Leu Lys Arg Tyr Ile Gly Phe Leu Leu Ile	
10 15 20	
TCT TTA GCG TTA TTA ATC ACG CCC TTT GTT CGC ATT GAT GGG GCG CAT	210
Ser Leu Ala Leu Leu Ile Thr Pro Phe Val Arg Ile Asp Gly Ala His	
25 30 35	
TTG TTT TTG ATC TCT TTT GAG CAT AAG CAA CTG CAT TTT TTA GGC AAG	258
Leu Phe Leu Ile Ser Phe Glu His Lys Gln Leu His Phe Leu Gly Lys	
40 45 50 55	
ATC TTT AGC GCT GAA GAA TTG CAA GTC ATG CCT TTT ATG GTT ATT TTG	306
Ile Phe Ser Ala Glu Glu Leu Gln Val Met Pro Phe Met Val Ile Leu	

-151-

																60	65
																70	
CTT	TTT	ATA	GGG	ATT	TTT	TTC	ATC	ACC	ACT	AGC	CTT	GGG	CGT	GTG	TGG	354	
Leu	Phe	Ile	Gly	Ile	Phe	Phe	Ile	Thr	Thr	Ser	Leu	Gly	Arg	Val	Trp		
75				80				85									
TGC	GGT	TGG	GCT	TGC	CCG	CAA	ACC	TTT	TTA	AGG	GTG	CTT	TAT	AGA	GAT	402	
Cys	Gly	Trp	Ala	Cys	Pro	Gln	Thr	Phe	Leu	Arg	Val	Leu	Tyr	Arg	Asp		
90				95				100									
GTG	ATT	GAA	ACC	AAG	ATT	TTC	AAA	CTC	CAT	AAA	AAG	ATC	AGC	AAC	AAG	450	
Val	Ile	Glu	Thr	Lys	Ile	Phe	Lys	Leu	His	Lys	Lys	Ile	Ser	Asn	Lys		
105				110				115									
CAA	GAA	AGC	CCT	AAA	AAC	ACC	CCA	AGC	TAC	AAG	ATC	CGT	AAA	GTA	TTG	498	
Gln	Glu	Ser	Pro	Lys	Asn	Thr	Pro	Ser	Tyr	Lys	Ile	Arg	Lys	Val	Leu		
120				125				130				135					
AGC	GTT	TTA	TTG	TTC	GCT	CCT	GTT	GTG	GCG	GGG	CTA	ATG	ATG	TTG	TTT	546	
Ser	Val	Leu	Leu	Phe	Ala	Pro	Val	Val	Ala	Gly	Leu	Met	Met	Leu	Phe		
140				145				150									
TTC	TTT	TAT	TTC	ATC	GCC	CCA	GAA	GAT	TTT	TTT	ATG	TAT	CTT	AAA	AAC	594	
Phe	Phe	Tyr	Phe	Ile	Ala	Pro	Glu	Asp	Phe	Phe	Met	Tyr	Leu	Lys	Asn		
155				160				165									
CCT	AGC	GAT	CAC	CCT	ATT	GCT	ATG	GGT	TTT	TGG	CTT	TTT	AGC	ACG	GCT	642	
Pro	Ser	Asp	His	Pro	Ile	Ala	Met	Gly	Phe	Trp	Leu	Phe	Ser	Thr	Ala		
170				175				180									
GTG	GTG	CTA	TTT	GAT	ATA	GTG	GTG	GTT	GCG	GAG	CGT	TTT	TGC	ATT	TAT	690	
Val	Val	Leu	Phe	Asp	Ile	Val	Val	Val	Ala	Glu	Arg	Phe	Cys	Ile	Tyr		
185				190				195									
TTA	TGC	CCT	TAC	GCT	AGG	GTG	CAA	TCG	GTG	TTG	TAT	GAC	AAT	GAC	ACC	738	
Leu	Cys	Pro	Tyr	Ala	Arg	Val	Gln	Ser	Val	Leu	Tyr	Asp	Asn	Asp	Thr		
200				205				210				215					
TTA	AAC	CCT	ATT	TAT	GAT	GAA	AAG	CGC	GGC	GGA	GCG	CTT	TAT	AAT	AAT	786	
Leu	Asn	Pro	Ile	Tyr	Asp	Glu	Lys	Arg	Gly	Gly	Ala	Leu	Tyr	Asn	Asn		
220				225				230									
CAG	GGC	CAT	CTC	TTC	CCC	TTA	CCT	CCC	AAA	AAA	CGC	AGC	CCA	GAA	AAC	834	
Gln	Gly	His	Leu	Phe	Pro	Leu	Pro	Pro	Lys	Lys	Arg	Ser	Pro	Glu	Asn		
235				240				245									
GAA	TGC	GTG	AAT	TGT	TTG	CAT	TGC	GTG	CAG	GTT	TGC	CCC	ACG	CAT	ATT	882	
Glu	Cys	Val	Asn	Cys	Leu	His	Cys	Val	Gln	Val	Cys	Pro	Thr	His	Ile		
250				255				260									
GAC	ATC	AGG	AAG	GGC	TTG	CAA	TTA	GAA	TGC	ATC	AAT	TGT	TTA	GAA	TGC	930	
Asp	Ile	Arg	Lys	Gly	Leu	Gln	Leu	Glu	Cys	Ile	Asn	Cys	Leu	Glu	Cys		
265				270				275									
GTG	GAT	GCA	TGC	ACG	ATT	ACC	ATG	GCT	AAA	TTT	AAC	CGC	CCT	TCA	CTC	978	

-152-

Val	Asp	Ala	Cys	Thr	Ile	Thr	Met	Ala	Lys	Phe	Asn	Arg	Pro	Ser	Leu		
280					285					290					295		
ATC	CAA	TGG	TCT	TCA	ACT	AAC	GCT	ATT	AAT	ACG	CGC	CAA	AAA	GTG	CAC	1026	
Ile	Gln	Trp	Ser	Ser	Thr	Asn	Ala	Ile	Asn	Thr	Arg	Gln	Lys	Val	His		
			300						305					310			
CTG	GTG	CGT	TTA	AAA	ACG	ATC	GCT	TAC	ATG	GGG	GTT	ATC	GCT	ATT	GTG	1074	
Leu	Val	Arg	Leu	Lys	Thr	Ile	Ala	Tyr	Met	Gly	Val	Ile	Ala	Ile	Val		
			315					320					325				
ATC	GCT	CTT	TTA	GCC	ATC	ACT	TCG	TTT	AAA	AAA	GAA	CGC	ATG	CTC	TTA	1122	
Ile	Ala	Leu	Leu	Ala	Ile	Thr	Ser	Phe	Lys	Lys	Glu	Arg	Met	Leu	Leu		
		330					335					340					
GAC	ATT	AAC	CGC	AAC	AGC	GAT	CTG	TAT	GAA	TTG	CGC	TCT	AGC	GGG	TAT	1170	
Asp	Ile	Asn	Arg	Asn	Ser	Asp	Leu	Tyr	Glu	Leu	Arg	Ser	Ser	Gly	Tyr		
	345					350					355						
GTG	GAT	AAC	GAT	TAC	GTG	TTT	TTA	TTC	CAC	AAC	ACG	GAC	AAT	AAA	GAC	1218	
Val	Asp	Asn	Asp	Tyr	Val	Phe	Leu	Phe	His	Asn	Thr	Asp	Asn	Lys	Asp		
360					365					370				375			
CAT	GAG	TTT	TAT	TTC	AAA	GTT	TTA	GGG	CAA	AAA	GAC	ATT	CAG	ATC	AAA	1266	
His	Glu	Phe	Tyr	Phe	Lys	Val	Leu	Gly	Gln	Lys	Asp	Ile	Gln	Ile	Lys		
			380					385					390				
AAG	CCT	TTA	AAT	CCT	ATC	GCC	ATT	AAA	GCC	GGG	CAA	AAG	ATT	AAA	GCG	1314	
Lys	Pro	Leu	Asn	Pro	Ile	Ala	Ile	Lys	Ala	Gly	Gln	Lys	Ile	Lys	Ala		
		395						400					405				
GTA	GTG	ATT	TTA	AGA	AAA	CCC	CTA	AAG	AGT	AAC	GCC	ACA	GAA	TAC	AAG	1362	
Val	Val	Ile	Leu	Arg	Lys	Pro	Leu	Lys	Ser	Asn	Ala	Thr	Glu	Tyr	Lys		
		410					415					420					
AAC	GCT	AAA	GAC	GCT	CTA	ATC	CCC	ATT	ACC	ATA	CAA	GCT	TAT	AGC	GCG	1410	
Asn	Ala	Lys	Asp	Ala	Leu	Ile	Pro	Ile	Thr	Ile	Gln	Ala	Tyr	Ser	Ala		
	425					430					435						
GAC	GAT	AAG	AAT	ATT	ACG	ATA	GAA	AGG	GAA	TCG	GTG	TTT	ATT	GCA	CCA	1458	
Asp	Asp	Lys	Asn	Ile	Thr	Ile	Glu	Arg	Glu	Ser	Val	Phe	Ile	Ala	Pro		
440					445					450				455			
AGT	GAG	GAT	TGAAGCCTAA	AACTAGCGTT	CAATCACTTC	ATAAGGCAAG	CCTTGTT									1514	
Ser	Glu	Asp															

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 458 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

-153-

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met	Leu	Glu	Thr	Ser	Ser	His	Phe	Leu	Lys	Ser	Phe	Arg	Leu	Lys	Arg	1	5	10	15
Tyr	Ile	Gly	Phe	Leu	Leu	Ile	Ser	Leu	Ala	Leu	Leu	Ile	Thr	Pro	Phe	20	25	30	
Val	Arg	Ile	Asp	Gly	Ala	His	Leu	Phe	Leu	Ile	Ser	Phe	Glu	His	Lys	35	40	45	
Gln	Leu	His	Phe	Leu	Gly	Lys	Ile	Phe	Ser	Ala	Glu	Glu	Leu	Gln	Val	50	55	60	
Met	Pro	Phe	Met	Val	Ile	Leu	Leu	Phe	Ile	Gly	Ile	Phe	Phe	Ile	Thr	65	70	75	80
Thr	Ser	Leu	Gly	Arg	Val	Trp	Cys	Gly	Trp	Ala	Cys	Pro	Gln	Thr	Phe	85	90	95	
Leu	Arg	Val	Leu	Tyr	Arg	Asp	Val	Ile	Glu	Thr	Lys	Ile	Phe	Lys	Leu	100	105	110	
His	Lys	Lys	Ile	Ser	Asn	Lys	Gln	Glu	Ser	Pro	Lys	Asn	Thr	Pro	Ser	115	120	125	
Tyr	Lys	Ile	Arg	Lys	Val	Leu	Ser	Val	Leu	Leu	Phe	Ala	Pro	Val	Val	130	135	140	
Ala	Gly	Leu	Met	Met	Leu	Phe	Phe	Phe	Tyr	Phe	Ile	Ala	Pro	Glu	Asp	145	150	155	160
Phe	Phe	Met	Tyr	Leu	Lys	Asn	Pro	Ser	Asp	His	Pro	Ile	Ala	Met	Gly	165	170	175	
Phe	Trp	Leu	Phe	Ser	Thr	Ala	Val	Val	Leu	Phe	Asp	Ile	Val	Val	Val	180	185	190	
Ala	Glu	Arg	Phe	Cys	Ile	Tyr	Leu	Cys	Pro	Tyr	Ala	Arg	Val	Gln	Ser	195	200	205	
Val	Leu	Tyr	Asp	Asn	Asp	Thr	Leu	Asn	Pro	Ile	Tyr	Asp	Glu	Lys	Arg	210	215	220	
Gly	Gly	Ala	Leu	Tyr	Asn	Asn	Gln	Gly	His	Leu	Phe	Pro	Leu	Pro	Pro	225	230	235	240
Lys	Lys	Arg	Ser	Pro	Glu	Asn	Glu	Cys	Val	Asn	Cys	Leu	His	Cys	Val	245	250	255	
Gln	Val	Cys	Pro	Thr	His	Ile	Asp	Ile	Arg	Lys	Gly	Leu	Gln	Leu	Glu	260	265	270	
Cys	Ile	Asn	Cys	Leu	Glu	Cys	Val	Asp	Ala	Cys	Thr	Ile	Thr	Met	Ala	275	280	285	
Lys	Phe	Asn	Arg	Pro	Ser	Leu	Ile	Gln	Trp	Ser	Ser	Thr	Asn	Ala	Ile	290	295	300	
Asn	Thr	Arg	Gln	Lys	Val	His	Leu	Val	Arg	Leu	Lys	Thr	Ile	Ala	Tyr	305	310	315	320
Met	Gly	Val	Ile	Ala	Ile	Val	Ile	Ala	Leu	Leu	Ala	Ile	Thr	Ser	Phe	325	330	335	
Lys	Lys	Glu	Arg	Met	Leu	Leu	Asp	Ile	Asn	Arg	Asn	Ser	Asp	Leu	Tyr	340	345	350	
Glu	Leu	Arg	Ser	Ser	Gly	Tyr	Val	Asp	Asn	Asp	Tyr	Val	Phe	Leu	Phe	355	360	365	
His	Asn	Thr	Asp	Asn	Lys	Asp	His	Glu	Phe	Tyr	Phe	Lys	Val	Leu	Gly	370	375	380	
Gln	Lys	Asp	Ile	Gln	Ile	Lys	Lys	Pro	Leu	Asn	Pro	Ile	Ala	Ile	Lys	385	390	395	400

-154-

```

Ala Gly Gln Lys Ile Lys Ala Val Val Ile Leu Arg Lys Pro Leu Lys
                405                      410                      415
Ser Asn Ala Thr Glu Tyr Lys Asn Ala Lys Asp Ala Leu Ile Pro Ile
                420                      425                      430
Thr Ile Gln Ala Tyr Ser Ala Asp Asp Lys Asn Ile Thr Ile Glu Arg
                435                      440                      445
Glu Ser Val Phe Ile Ala Pro Ser Glu Asp
                450                      455

```

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 228...782
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```

ACGATTGAT CAATAACGAA AATAAAATTG ATGAAATCAA TAATGAAGAA AACGCTGATC      60
CTTCGCAAAA AAGAACGAAC AACGTTTTGC AACGAGCCAC TAACCACCAA GACAATCTCA      120
ATTCCCCACT CAACAGGAAG TATTAAAGTG TGAAACTTTT TTCAAAGGAT TTATTTAAAA      180
AAGTAACCCC TTTATTTTTA AGCGTTTATT TTTTAAACCC CACCATT ATG CAA GCC      236
                                   Met Gln Ala
                                   1

AAA AGC CGT TTT TAT GTG GCT TCT CAA TAC CAG GTG GGG AAA ATG ATC      284
Lys Ser Arg Phe Tyr Val Ala Ser Gln Tyr Gln Val Gly Lys Met Ile
    5                      10                      15

ATG AAA AAA TAC AAC GAT CTC AAA CGC ACG ATT GAA GGG GCG AGC TTT      332
Met Lys Lys Tyr Asn Asp Leu Lys Arg Thr Ile Glu Gly Ala Ser Phe
    20                      25                      30                      35

TCT TTA GGC TGG GAG ATT AAC CCC ACT AAC TAC TGG TTT TAT TCG CGC      380
Ser Leu Gly Trp Glu Ile Asn Pro Thr Asn Tyr Trp Phe Tyr Ser Arg
                40                      45                      50

TAT TAC TTT TTT ATG GAT TAC GGG AAT GTC ATT CTC AAT AAA AGA ACG      428
Tyr Tyr Phe Phe Met Asp Tyr Gly Asn Val Ile Leu Asn Lys Arg Thr
                55                      60                      65

GGC GCT CAA GCG AAC ATG TTC ACT TAT GGC TTT GGG GGG GAT TTG ATT      476
Gly Ala Gln Ala Asn Met Phe Thr Tyr Gly Phe Gly Gly Asp Leu Ile
                70                      75                      80

GTG GAA TAC AAT AAA AAC CCC TTG TAT GTA TTT TCT CTT TTT TAT GGC      524

```

-155-

Val	Glu	Tyr	Asn	Lys	Asn	Pro	Leu	Tyr	Val	Phe	Ser	Leu	Phe	Tyr	Gly	
85						90				95						
ATG	CAA	GTT	GCT	GAA	AAC	ACA	TGG	ACG	ATT	TCC	AAA	CAC	AGC	GCG	AAT	572
Met	Gln	Val	Ala	Glu	Asn	Thr	Trp	Thr	Ile	Ser	Lys	His	Ser	Ala	Asn	
100					105					110					115	
TTC	ATC	ATT	GAC	GAT	TGG	CGC	AGC	ATT	CAA	GGG	TTT	TCG	CTC	AAA	ACT	620
Phe	Ile	Ile	Asp	Asp	Trp	Arg	Ser	Ile	Gln	Gly	Phe	Ser	Leu	Lys	Thr	
				120					125					130		
TCC	AAT	TTT	AGG	ATG	TTG	GGT	TTA	GTG	GGG	TTT	AAA	TTC	CAA	ACC	GTG	668
Ser	Asn	Phe	Arg	Met	Leu	Gly	Leu	Val	Gly	Phe	Lys	Phe	Gln	Thr	Val	
			135					140					145			
CTA	TTC	CAC	CAT	GAC	GCA	AGT	ATT	GAA	GTG	GGG	ATC	AAA	TGG	CCT	TTT	716
Leu	Phe	His	His	Asp	Ala	Ser	Ile	Glu	Val	Gly	Ile	Lys	Trp	Pro	Phe	
		150					155					160				
GCT	TTT	GAA	TAC	GAC	TCA	GCC	TTT	GTA	AGG	CTT	TTT	TCT	GTC	TTT	ATT	764
Ala	Phe	Glu	Tyr	Asp	Ser	Ala	Phe	Val	Arg	Leu	Phe	Ser	Val	Phe	Ile	
	165					170				175						
TCG	CAC	ACT	TTC	TAC	CTT	TAACTAATT	CCAACCCTAC	CGGGCAATGA	TCGCTCCC							820
Ser	His	Thr	Phe	Tyr	Leu											
180					185											
TAAAATATCT	TTATAGATTA	AAGCGTCTTT	TAAGCGCGTT	TTTAAAGGGT	TAGAGCATAA											880
AAAATAATCA	ATGCGCCAAC	CAATGTTTTT	ATCCCTTGCT	TGTTGCATGT	AACTCCACCA											940
GGTGTAAGCC	TTTTCTTTGT	TAGGGTAAAA	ATAACGGA	GTGTCAATAA												990

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 185 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Met	Gln	Ala	Lys	Ser	Arg	Phe	Tyr	Val	Ala	Ser	Gln	Tyr	Gln	Val	Gly	
1				5					10					15		
Lys	Met	Ile	Met	Lys	Lys	Tyr	Asn	Asp	Leu	Lys	Arg	Thr	Ile	Glu	Gly	
			20					25					30			
Ala	Ser	Phe	Ser	Leu	Gly	Trp	Glu	Ile	Asn	Pro	Thr	Asn	Tyr	Trp	Phe	
		35				40					45					
Tyr	Ser	Arg	Tyr	Tyr	Phe	Phe	Met	Asp	Tyr	Gly	Asn	Val	Ile	Leu	Asn	
	50				55				60							
Lys	Arg	Thr	Gly	Ala	Gln	Ala	Asn	Met	Phe	Thr	Tyr	Gly	Phe	Gly	Gly	
65					70				75					80		
Asp	Leu	Ile	Val	Glu	Tyr	Asn	Lys	Asn	Pro	Leu	Tyr	Val	Phe	Ser	Leu	
			85					90						95		

-156-

```

Phe Tyr Gly Met Gln Val Ala Glu Asn Thr Trp Thr Ile Ser Lys His
      100                      105                      110
Ser Ala Asn Phe Ile Ile Asp Asp Trp Arg Ser Ile Gln Gly Phe Ser
      115                      120                      125
Leu Lys Thr Ser Asn Phe Arg Met Leu Gly Leu Val Gly Phe Lys Phe
      130                      135                      140
Gln Thr Val Leu Phe His His Asp Ala Ser Ile Glu Val Gly Ile Lys
      145                      150                      155                      160
Trp Pro Phe Ala Phe Glu Tyr Asp Ser Ala Phe Val Arg Leu Phe Ser
      165                      170                      175
Val Phe Ile Ser His Thr Phe Tyr Leu
      180                      185

```

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1161 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 109...1113
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```

ATCTTACCTT TATCTTTTAA GATTTTATGA AAAATAGTTT CATTTTTTACT ATTGTTATTT      60
TCTTAGTAAT GTTATAATCG CTTTATAAAT CATACAAAAA GGATCGCT ATG TTA GTT      117
                                     Met Leu Val
                                     1

ACT CGC TTT AAA AAA GCT TTC ATT TCT TAT TCT TTA GGC GTG CTT GTC      165
Thr Arg Phe Lys Lys Ala Phe Ile Ser Tyr Ser Leu Gly Val Leu Val
      5                      10                      15

GCT TCA TTA TGG TTG AAC GTG TGC AAC GCT TCA GCG CAA GAA GTC AAA      213
Ala Ser Leu Trp Leu Asn Val Cys Asn Ala Ser Ala Gln Glu Val Lys
      20                      25                      30                      35

GTC AAG GAT TAT TTC GGG GAG CAA ACC ATC AAG CTT CCT GTT TCT AAA      261
Val Lys Asp Tyr Phe Gly Glu Gln Thr Ile Lys Leu Pro Val Ser Lys
      40                      45                      50

ATA GCC TAT ATA GGG AGC TAT GTA GAA GTG CCT GCC ATG CTT AAT GTT      309
Ile Ala Tyr Ile Gly Ser Tyr Val Glu Val Pro Ala Met Leu Asn Val
      55                      60                      65

TGG AAT AGG GTT GTA GGC GTT TCG GAT TAC GCT TTT AAA GAC GAT ATT      357
Trp Asn Arg Val Val Gly Val Ser Asp Tyr Ala Phe Lys Asp Asp Ile
      70                      75                      80

```

-157-

GTC AAA GCC ACT CTC AAA GGC GAA GAT CTT AAA CGC GTC AAA CAC ATG Val Lys Ala Thr Leu Lys Gly Glu Asp Leu Lys Arg Val Lys His Met 85 90 95	405
AGC ACT GAT CAT ACA GCC GCG CTA AAT GTA GAG CTT TTA AAA AAG CTT Ser Thr Asp His Thr Ala Ala Leu Asn Val Glu Leu Leu Lys Lys Leu 100 105 110 115	453
AGC CCT GAT CTT GTG GTA ACC TTT GTG GGC AAC CCT AAA GCG GTA GAG Ser Pro Asp Leu Val Val Thr Phe Val Gly Asn Pro Lys Ala Val Glu 120 125 130	501
CAT GCG AAA AAA TTT GGT ATA TCA TTT CTT TCT TTT CAA GAG ACA ACG His Ala Lys Lys Phe Gly Ile Ser Phe Leu Ser Phe Gln Glu Thr Thr 135 140 145	549
ATT GCA GAG GCC ATG CAG GCC ATG CAA GCT CAA GCC ACG GTT TTA GAG Ile Ala Glu Ala Met Gln Ala Met Gln Ala Gln Ala Thr Val Leu Glu 150 155 160	597
ATT GAC GCT TCC AAA AAA TTC GCC AAA ATG CAA GAA ACT TTG GAT TTT Ile Asp Ala Ser Lys Lys Phe Ala Lys Met Gln Glu Thr Leu Asp Phe 165 170 175	645
ATT GCT GAG CGT TTG AAA AAT GTC AAA AAG AAA AAG GGG GTG GAG CTT Ile Ala Glu Arg Leu Lys Asn Val Lys Lys Lys Lys Gly Val Glu Leu 180 185 190 195	693
TTC CAT AAA GCC AAT AAA ATC AGC GGC CAT CAA GCC ATT AGC TCA GAC Phe His Lys Ala Asn Lys Ile Ser Gly His Gln Ala Ile Ser Ser Asp 200 205 210	741
ATT TTA GAA AAA GGG GGC ATA GAC AAT TTT GGC TTG AAA TAT GTC AAA Ile Leu Glu Lys Gly Gly Ile Asp Asn Phe Gly Leu Lys Tyr Val Lys 215 220 225	789
TTT GGG CGT GCT GAC ATT AGC GTG GAA AAA ATC GTT AAA GAA AAC CCT Phe Gly Arg Ala Asp Ile Ser Val Glu Lys Ile Val Lys Glu Asn Pro 230 235 240	837
GAG ATT ATC TTT ATT TGG TGG ATA AGC CCA CTC ACG CCT GAA GAT GTG Glu Ile Ile Phe Ile Trp Trp Ile Ser Pro Leu Thr Pro Glu Asp Val 245 250 255	885
TTA AAC AAC CCC AAA TTT GCT ACC ATC AAA GCC ATT AAA AAC AAG CAG Leu Asn Asn Pro Lys Phe Ala Thr Ile Lys Ala Ile Lys Asn Lys Gln 260 265 270 275	933
GTT TAT AAA CTC CCC ACA ATG GAT ATT GGC GGG CCT AGA GCC CCA CTC Val Tyr Lys Leu Pro Thr Met Asp Ile Gly Gly Pro Arg Ala Pro Leu 280 285 290	981
ATA AGT CTT TTT ATC GCT CTA AAA GCC CAC CCT GAA GCC TTT AAG GGC Ile Ser Leu Phe Ile Ala Leu Lys Ala His Pro Glu Ala Phe Lys Gly 295 300 305	1029

GTG	GAT	ATT	AAT	GCG	ATG	GTT	AAA	GAC	TAC	TAT	AAA	GTG	GTT	TTT	GAT	1077
Val	Asp	Ile	Asn	Ala	Met	Val	Lys	Asp	Tyr	Tyr	Lys	Val	Val	Phe	Asp	
		310					315					320				
TTG	AAT	GAT	GCA	GAG	GTT	GAG	CCC	TTT	TTA	TGG	CAT	TAATTTT	TAA	AAAGGG	1129	
Leu	Asn	Asp	Ala	Glu	Val	Glu	Pro	Phe	Leu	Trp	His					
		325				330					335					
GTG	GAT	ATT	AAT	GCG	ATG	GTT	AAA	GAC	TAC	TAT	AAA	GTG	GTT	TTT	GAT	1161
Val	Asp	Ile	Asn	Ala	Met	Val	Lys	Asp	Tyr	Tyr	Lys	Val	Val	Phe	Asp	
		310					315					320				

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

BNSDOCID: <WO_9821225A1_I_>

-159-

Glu	Asp	Val	Leu	Asn	Asn	Pro	Lys	Phe	Ala	Thr	Ile	Lys	Ala	Ile	Lys
			260					265					270		
Asn	Lys	Gln	Val	Tyr	Lys	Leu	Pro	Thr	Met	Asp	Ile	Gly	Gly	Pro	Arg
		275					280					285			
Ala	Pro	Leu	Ile	Ser	Leu	Phe	Ile	Ala	Leu	Lys	Ala	His	Pro	Glu	Ala
	290					295					300				
Phe	Lys	Gly	Val	Asp	Ile	Asn	Ala	Met	Val	Lys	Asp	Tyr	Tyr	Lys	Val
305					310					315				320	
Val	Phe	Asp	Leu	Asn	Asp	Ala	Glu	Val	Glu	Pro	Phe	Leu	Trp	His	
			325					330						335	

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 800 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 121...669
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTATTTCGCAT GCATTAGCTA TTATTGAAGC TCAAAGCATT CAAGCGCATT TATTCTTAGA	60
TGAAATCAAA CAAAGCCAAA AAGAAAAGAA AAAATTCCCC ACTTTCAAAG GAGGTTTTTA	120
ATG CGT TGG TGG TGT TTT TTG GTG TGT TGT TTT GGT ATT TTA AGC GTG	168
Met Arg Trp Trp Cys Phe Leu Val Cys Cys Phe Gly Ile Leu Ser Val	
1 5 10 15	
ATG GAC GCT AAA AAA TTA GAG AAT AAG AAT TTG AAA AAA GAA AGA GAG	216
Met Asp Ala Lys Lys Leu Glu Asn Lys Asn Leu Lys Lys Glu Arg Glu	
20 25 30	
CTT TTA GAG ATT ACT GGC AAC CAA TTT GTA GCG AAC GAC AAA ACC AAA	264
Leu Leu Glu Ile Thr Gly Asn Gln Phe Val Ala Asn Asp Lys Thr Lys	
35 40 45	
ACC GCT GTT ATT CAA GGC AAT GTG CAG ATC AAA AAG GGT AAA GAC CGG	312
Thr Ala Val Ile Gln Gly Asn Val Gln Ile Lys Lys Gly Lys Asp Arg	
50 55 60	
TTG TTT GCG GAC AAG GTG AGC GTG TTT TTA AAC GAT AAA CGA AAG CCA	360
Leu Phe Ala Asp Lys Val Ser Val Phe Leu Asn Asp Lys Arg Lys Pro	
65 70 75 80	
GAG CGC TAT GAA GCC ACA GGG AAC ACG CAT TTT AAC ATC TTT ACA GAG	408
Glu Arg Tyr Glu Ala Thr Gly Asn Thr His Phe Asn Ile Phe Thr Glu	
85 90 95	

-160-

```

GAC AAT CGT GAA ATC AGC GGG AGT GCT GAC AAG CTC ATT TAT AAC GCG      456
Asp Asn Arg  Glu Ile Ser Gly Ser Ala Asp Lys Leu Ile Tyr Asn Ala
      100                      105                      110

CTG AAT GGG GAA TAC AAA TTA TTG CAA AAT GCG GTG GTT AGA GAA GTG      504
Leu Asn Gly Glu Tyr Lys Leu Leu Gln Asn Ala Val Val Arg Glu Val
      115                      120                      125

GGG AAA TCC AAT GTC ATC ACC GGC GAT GAA ATC ATT TTA AAC AAA ACT      552
Gly Lys Ser Asn Val Ile Thr Gly Asp Glu Ile Ile Leu Asn Lys Thr
      130                      135                      140

AAG GGT TAT GCT GAT GTG TTG GGG AGC GCG AAA CGG CCC GCT AAA TTC      600
Lys Gly Tyr Ala Asp Val Leu Gly Ser Ala Lys Arg Pro Ala Lys Phe
      145                      150                      155                      160

GTG TTT GAT ATG GAA GAT ATT AAT GAA GAA AAT CGT AAG GCT AAA TTG      648
Val Phe Asp Met Glu Asp Ile Asn Glu Glu Asn Arg Lys Ala Lys Leu
      165                      170                      175

AAG AAG AAA GGC GAA AAA CCA TGATTGTCAT TAAAGACGCT CATTTTCTCA CTTC      703
Lys Lys Lys Gly Glu Lys Pro
      180

TTCAAGCCAA CTTTTTCAAT GCCCTGCGAG TTTGACTTCT GAAATGGTGG TTTTAGGGCG      763
CAGCAATGTA GGCAAAAGCT CGTTTATTAA TACCTTG
      800

```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

Met Arg Trp Trp Cys Phe Leu Val Cys Cys Phe Gly Ile Leu Ser Val
 1           5           10           15
Met Asp Ala Lys Lys Leu Glu Asn Lys Asn Leu Lys Lys Glu Arg Glu
      20           25           30
Leu Leu Glu Ile Thr Gly Asn Gln Phe Val Ala Asn Asp Lys Thr Lys
      35           40           45
Thr Ala Val Ile Gln Gly Asn Val Gln Ile Lys Lys Gly Lys Asp Arg
      50           55           60
Leu Phe Ala Asp Lys Val Ser Val Phe Leu Asn Asp Lys Arg Lys Pro
      65           70           75           80
Glu Arg Tyr Glu Ala Thr Gly Asn Thr His Phe Asn Ile Phe Thr Glu
      85           90           95
Asp Asn Arg Glu Ile Ser Gly Ser Ala Asp Lys Leu Ile Tyr Asn Ala
      100          105          110
Leu Asn Gly Glu Tyr Lys Leu Leu Gln Asn Ala Val Val Arg Glu Val

```

-161-

115	120	125
Gly Lys Ser Asn Val Ile Thr Gly Asp Glu Ile Ile Leu Asn Lys Thr		
130	135	140
Lys Gly Tyr Ala Asp Val Leu Gly Ser Ala Lys Arg Pro Ala Lys Phe		
145	150	155
Val Phe Asp Met Glu Asp Ile Asn Glu Glu Asn Arg Lys Ala Lys Leu		
165	170	175
Lys Lys Lys Gly Glu Lys Pro		
180		

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 724 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 88...618
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GTGATGATTG AAAACATGTG AAAGAGCGTT TTTTAAGCTT TTAAATGGTG TTTGAATGCG	60
AAAAAAAGGC TAATACTATC ATAAGGA ATG AAG TTG ATA AAA TTT GTG CGT AAT	114
Met Lys Leu Ile Lys Phe Val Arg Asn	
1 5	
GTG GTT TTG TTC ATT TTA ACG GCG ATC TTT TTA GCG TTC ATG CTT TTG	162
Val Val Leu Phe Ile Leu Thr Ala Ile Phe Leu Ala Phe Met Leu Leu	
10 15 20 25	
GTG AGT TAT TGC ATG CCC CAT TAT AGC GCG GCT GTC ATT AGC GGG GTG	210
Val Ser Tyr Cys Met Pro His Tyr Ser Ala Ala Val Ile Ser Gly Val	
30 35 40	
GAA GTC AAA AGA ATG AAT GAA AAT GAA AAC ACG CCC AAT AAT AAG GAA	258
Glu Val Lys Arg Met Asn Glu Asn Glu Asn Thr Pro Asn Asn Lys Glu	
45 50 55	
GTA AAA ACC CTT GCT AGA GAT GTC TAT TTT GTG CAA ACT TAC GAC CCT	306
Val Lys Thr Leu Ala Arg Asp Val Tyr Phe Val Gln Thr Tyr Asp Pro	
60 65 70	
AAA GAT CAA AAA AGC GTA ACC GTT TAT CGT AAC GAA GAC ACG CGC TTT	354
Lys Asp Gln Lys Ser Val Thr Val Tyr Arg Asn Glu Asp Thr Arg Phe	
75 80 85	
AGC TTC CCT TTT TAT TTT AAG TTT AAT TCG GCT GAT ATT TCA GCC CTC	402
Ser Phe Pro Phe Tyr Phe Lys Phe Asn Ser Ala Asp Ile Ser Ala Leu	

-162-

90	95	100	105	
GCT CAA AGT TTA ATC AAT CAG CAA GTG GAA GTG AAA TAC TAT GGT TGG				450
Ala Gln Ser Leu Ile Asn Gln Gln Val Glu Val Lys Tyr Tyr Gly Trp				
110		115	120	
CGG ATC AAT TTG TTT AAC ATG TTC CCT AAT GTG ATT TTT TTA AAG CCC				498
Arg Ile Asn Leu Phe Asn Met Phe Pro Asn Val Ile Phe Leu Lys Pro				
125	130		135	
TTA AAA GAG AGC ACT GAC ATT TCA AAG CCC ATT TTT AGC TGG ATT TTA				546
Leu Lys Glu Ser Thr Asp Ile Ser Lys Pro Ile Phe Ser Trp Ile Leu				
140	145		150	
TAC GCT TTG CTG TTA ATG GGC TTT TTT ATC AGC GCG CGT TCT GTT TGC				594
Tyr Ala Leu Leu Leu Met Gly Phe Phe Ile Ser Ala Arg Ser Val Cys				
155	160		165	
ACT TTA TTT AAG AGC AAA GCT CAT TAAACTTTT AGGCTTTGTT GGAAAATCAC				648
Thr Leu Phe Lys Ser Lys Ala His				
170	175			
AATGGGGTTA TTGGAGCGTG TATTAAAAAG CTCAATATAG GGCAAGCTGA TGCTGTGAAA				708
AGCGGTGTTG TTTCTT				724

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 177 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Met	Lys	Leu	Ile	Lys	Phe	Val	Arg	Asn	Val	Val	Leu	Phe	Ile	Leu	Thr
1				5					10					15	
Ala	Ile	Phe	Leu	Ala	Phe	Met	Leu	Leu	Val	Ser	Tyr	Cys	Met	Pro	His
			20					25					30		
Tyr	Ser	Ala	Ala	Val	Ile	Ser	Gly	Val	Glu	Val	Lys	Arg	Met	Asn	Glu
		35					40					45			
Asn	Glu	Asn	Thr	Pro	Asn	Asn	Lys	Glu	Val	Lys	Thr	Leu	Ala	Arg	Asp
		50				55					60				
Val	Tyr	Phe	Val	Gln	Thr	Tyr	Asp	Pro	Lys	Asp	Gln	Lys	Ser	Val	Thr
65					70					75				80	
Val	Tyr	Arg	Asn	Glu	Asp	Thr	Arg	Phe	Ser	Phe	Pro	Phe	Tyr	Phe	Lys
			85						90					95	
Phe	Asn	Ser	Ala	Asp	Ile	Ser	Ala	Leu	Ala	Gln	Ser	Leu	Ile	Asn	Gln
			100					105					110		
Gln	Val	Glu	Val	Lys	Tyr	Tyr	Gly	Trp	Arg	Ile	Asn	Leu	Phe	Asn	Met
		115					120					125			
Phe	Pro	Asn	Val	Ile	Phe	Leu	Lys	Pro	Leu	Lys	Glu	Ser	Thr	Asp	Ile

-163-

130		135		140
Ser Lys Pro Ile Phe Ser Trp Ile Leu Tyr Ala Leu Leu Leu Met Gly				
145		150		155
Phe Phe Ile Ser Ala Arg Ser Val Cys Thr Leu Phe Lys Ser Lys Ala				
	165		170	175
His				

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 982 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 117...911
- (D) OTHER INFORMATION:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 117...167
- (D) OTHER INFORMATION:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 168...911
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCACATTTAA GGTAGAAACC ACTCAATTAG ATGTAAAAAT TCCAAACGGC AACCAAAAAA	60
TGGTTAAAAA GGACACAATA AACCCCAAAA ATGAAATTTA AATATATGGG AACTTA ATG	119
	Met
	-17
AGA ATT TTT TTT GTT ATC ATG GGA CTT GTG TTT TTT GGT TGC ACC AGT	167
Arg Ile Phe Phe Val Ile Met Gly Leu Val Phe Phe Gly Cys Thr Ser	
-15 -10 -5	
AAG GTG CAT GAG ATG AAA AAA AGC CCT TGC ACC TTG TAT GAA AAC AGG	215
Lys Val His Glu Met Lys Lys Ser Pro Cys Thr Leu Tyr Glu Asn Arg	
1 5 10 15	
TTA AAT CTC GCA GAA ATC TTT CAC AAG CGA GCA ATT GAT CTA TTT AGA	263
Leu Asn Leu Ala Glu Ile Phe His Lys Arg Ala Ile Asp Leu Phe Arg	
20 25 30	
GAG CTT TTA AGC CAC CAA GAA AAG CAT TTA GAA AAC AAG CTT TCT GGT	311
Glu Leu Leu Ser His Gln Glu Lys His Leu Glu Asn Lys Leu Ser Gly	

-164-

35	40	45	
TTT TCG GTG AGT GAT TTG GAC ATG CAA AGC GTG TTT CGG CTG GAA AGA			359
Phe Ser Val Ser Asp Leu Asp Met Gln Ser Val Phe Arg Leu Glu Arg			
50	55	60	
AAC CGC TTG AAA ATC GCT TAC AAG CTC TTA GGC TTG ATG AGT TTT ATC			407
Asn Arg Leu Lys Ile Ala Tyr Lys Leu Leu Gly Leu Met Ser Phe Ile			
65	70	75	80
GCT CTT ATT TTA GCG ATC GTG TTA ATC AGT CTT CTA CCC TTA CAA AAA			455
Ala Leu Ile Leu Ala Ile Val Leu Ile Ser Leu Leu Pro Leu Gln Lys			
85	90		95
ACC GAA CAC CAT TTC GTG GAT TTT TTA AAC CAG GAC AAG CAT TAC GTC			503
Thr Glu His His Phe Val Asp Phe Leu Asn Gln Asp Lys His Tyr Val			
100	105		110
ATT ATC CAA AGA GCG GAT AAA AGC ATT TCC AGT AAT GAA GCG TTG GCT			551
Ile Ile Gln Arg Ala Asp Lys Ser Ile Ser Ser Asn Glu Ala Leu Ala			
115	120		125
CGT TCG CTC ATT GGG GCG TAT GTG TTA AAC CGA GAG AGC ATT AAC CGC			599
Arg Ser Leu Ile Gly Ala Tyr Val Leu Asn Arg Glu Ser Ile Asn Arg			
130	135		140
ATT GAC GAT AAA TCG CGC TAT GAA TTG GTG CGC TTG CAA AGC AGT TCT			647
Ile Asp Asp Lys Ser Arg Tyr Glu Leu Val Arg Leu Gln Ser Ser Ser			
145	150	155	160
AAA GTG TGG CAA CGC TTT GAA GAT TTG ATT AAA ACC CAA AAC AGC ATT			695
Lys Val Trp Gln Arg Phe Glu Asp Leu Ile Lys Thr Gln Asn Ser Ile			
165	170		175
TAT GTG CAA AGC CAT TTG GAA AGA GAA GTC CAT ATC GTC AAT ATT GCG			743
Tyr Val Gln Ser His Leu Glu Arg Glu Val His Ile Val Asn Ile Ala			
180	185		190
ATC TAT CAG CAA GAC AAT AAC CCC ATT GCG AGC GTC TCC ATT GCC GCT			791
Ile Tyr Gln Gln Asp Asn Asn Pro Ile Ala Ser Val Ser Ile Ala Ala			
195	200		205
AAA CTT TTG AAT GAA AAC AAG CTG GTG TAT GAA AAG CGT TAT AAA ATC			839
Lys Leu Leu Asn Glu Asn Lys Leu Val Tyr Glu Lys Arg Tyr Lys Ile			
210	215		220
GTA TTG AGT TAT TTG TTT GAC ACC CCG ATG AAT TCA AGC TTG CAA GCT			887
Val Leu Ser Tyr Leu Phe Asp Thr Pro Met Asn Ser Ser Leu Gln Ala			
225	230	235	240
TGC AAG CTC TCA GGC TTC ATA GTT TGACATGACA TATAGATGAG CTTTATGCGG			941
Cys Lys Leu Ser Gly Phe Ile Val			
245			
TACGATTATC ACAGAATGGC TAACGCAGCA GGCACCGAGT A			982

-165-

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```

Met Arg Ile Phe Phe Val Ile Met Gly Leu Val Phe Phe Gly Cys Thr
-17   -15               -10               -5
Ser Lys Val His Glu Met Lys Lys Ser Pro Cys Thr Leu Tyr Glu Asn
  1             5             10             15
Arg Leu Asn Leu Ala Glu Ile Phe His Lys Arg Ala Ile Asp Leu Phe
                20             25             30
Arg Glu Leu Leu Ser His Gln Glu Lys His Leu Glu Asn Lys Leu Ser
                35             40             45
Gly Phe Ser Val Ser Asp Leu Asp Met Gln Ser Val Phe Arg Leu Glu
  50             55             60
Arg Asn Arg Leu Lys Ile Ala Tyr Lys Leu Leu Gly Leu Met Ser Phe
  65             70             75
Ile Ala Leu Ile Leu Ala Ile Val Leu Ile Ser Leu Leu Pro Leu Gln
  80             85             90             95
Lys Thr Glu His His Phe Val Asp Phe Leu Asn Gln Asp Lys His Tyr
                100            105            110
Val Ile Ile Gln Arg Ala Asp Lys Ser Ile Ser Ser Asn Glu Ala Leu
                115            120            125
Ala Arg Ser Leu Ile Gly Ala Tyr Val Leu Asn Arg Glu Ser Ile Asn
  130            135            140
Arg Ile Asp Asp Lys Ser Arg Tyr Glu Leu Val Arg Leu Gln Ser Ser
  145            150            155
Ser Lys Val Trp Gln Arg Phe Glu Asp Leu Ile Lys Thr Gln Asn Ser
  160            165            170            175
Ile Tyr Val Gln Ser His Leu Glu Arg Glu Val His Ile Val Asn Ile
                180            185            190
Ala Ile Tyr Gln Gln Asp Asn Asn Pro Ile Ala Ser Val Ser Ile Ala
  195            200            205
Ala Lys Leu Leu Asn Glu Asn Lys Leu Val Tyr Glu Lys Arg Tyr Lys
  210            215            220
Ile Val Leu Ser Tyr Leu Phe Asp Thr Pro Met Asn Ser Ser Leu Gln
  225            230            235
Ala Cys Lys Leu Ser Gly Phe Ile Val
  240            245

```

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2059 base pairs

-166-

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 183...1961
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GATATTTGTT TTGTTGGGGG TTAGGTTTTT GTTTAAGAAA GTTTTTTAAA ACTAAAGAAG	60
CGCTTAAAC AGAACCTTTT GTTTTTTAGG TTTTATTTT TACTTTGGCT TGTTTCAAA	120
AGTCATTTTG ATTTCTAAAA ATAGTCTATA ATGCTCGCAA GAGATATTTT TTAAGGTTAT	180
CA ATG AAA GCT ATA AAA ATA CTT TTT ATA ATG ACA CTC AGT TTA AAC	227
Met Lys Ala Ile Lys Ile Leu Phe Ile Met Thr Leu Ser Leu Asn	
1 5 10 15	
GCT ATC AGC GTG AAT AGG GCG TTG TTT GAT TTA AAA GAT TCG CAA TTA	275
Ala Ile Ser Val Asn Arg Ala Leu Phe Asp Leu Lys Asp Ser Gln Leu	
20 25 30	
AAA GGG GAA TTA ACG CCA AAA ATA GTG AAT TTT GGG GGT TAT AAA AGC	323
Lys Gly Glu Leu Thr Pro Lys Ile Val Asn Phe Gly Gly Tyr Lys Ser	
35 40 45	
AGC ACT GAA GAG TGG GGG GCT ACG GCT TTA AAC TAT ATC AAT GCG GCT	371
Ser Thr Glu Glu Trp Gly Ala Thr Ala Leu Asn Tyr Ile Asn Ala Ala	
50 55 60	
AAT GGC GAT GCG AAA AAA TTC AGC ACT CTA GTG GAA AAA ATG CGT TTT	419
Asn Gly Asp Ala Lys Lys Phe Ser Thr Leu Val Glu Lys Met Arg Phe	
65 70 75	
AAC TCC GGT ATA TTG GGG AAT TTA AGA GTG CAT GCA CGT TTG AGG CAA	467
Asn Ser Gly Ile Leu Gly Asn Leu Arg Val His Ala Arg Leu Arg Gln	
80 85 90 95	
GCC CTA AAA TTG CAA AAG AAT TTG AAA TAT TGC CTT AAA ATC ATC GCT	515
Ala Leu Lys Leu Gln Lys Asn Leu Lys Tyr Cys Leu Lys Ile Ile Ala	
100 105 110	
AGG GAT TCT TTT TAT AGC TAC CGC ACC GGT ATT TAT ATC CCC TTA GGC	563
Arg Asp Ser Phe Tyr Ser Tyr Arg Thr Gly Ile Tyr Ile Pro Leu Gly	
115 120 125	
ATT TCT TTA AAA GAT CAA AAA ACG GCT CAA AAA ATG CTC GCT GAT TTG	611
Ile Ser Leu Lys Asp Gln Lys Thr Ala Gln Lys Met Leu Ala Asp Leu	
130 135 140	
AGC GTG GTA GGG GCG TAT CTT AAA AAA CAA CAA GAG AAT GAA AAG GCT	659
Ser Val Val Gly Ala Tyr Leu Lys Lys Gln Gln Glu Asn Glu Lys Ala	

-167-

145	150	155	
CAA AGC CCT TAT TAC Gln Ser Pro Tyr Tyr 160	AGA AAC AAC AAC Arg Asn Asn Asn 165	TAT TAC AAC TCT TAC TAT AGC Tyr Tyr Asn Ser Tyr Tyr Ser 170	707
CCT TAT TAC GGA ATG Pro Tyr Tyr Gly Met 180	TAT GGT ATG TAT GGC ATG Tyr Gly Met Tyr Gly Met 185	GGC ATG TAT GGA ATG Gly Met Tyr Gly Met 190	755
TAT GGC ATG GGC ATG Tyr Gly Met Gly Met 195	TAT GAT TTT TAT GAC TTT Tyr Asp Phe Tyr Asp 200	TAT GAT GGC ATG TAT Tyr Asp Gly Met Tyr 205	803
GGA TTC TAC CCT AAC ATG Gly Phe Tyr Pro Asn Met 210	TTT TTC ATG ATG CAA GTT Phe Phe Met Met Gln Val 215	CAA GAT TAC TTG Gln Asp Tyr Leu 220	851
ATG TTA GAA AAT TAC ATG Met Leu Glu Asn Tyr Met 225	TAT GCG CTC GAT CAA GAA Tyr Ala Leu Asp Gln Glu 230	GAG ATT TTA GAT Glu Glu Ile Leu Asp 235	899
CAT GAC GCT TCT ACT His Asp Ala Ser Thr 240	GAC CAA CTT GAT ACG CCT Asp Gln Leu Asp Thr Pro 245	ACT GAT GAT GAC AAA Thr Asp Asp Asp Lys 250	947
GAC GAT AAA GAC GAT AAA Asp Asp Lys Asp Asp Lys 260	TCC TTA CAG CAG GCA Ser Leu Gln Gln Ala 265	AAT CTT ATG AAC TTT Asn Leu Met Asn Phe 270	995
TAT CGT GAT CCC AAA TTC Tyr Arg Asp Pro Lys Phe 275	AGC AAA GGC ATT CAA ACC Ser Lys Gly Ile Gln Thr 280	AAC CGC TTG AAT Asn Arg Leu Asn 285	1043
AGC GCT TTA GTC AAT TTA Ser Ala Leu Val Asn Leu 290	GAC AAC AGT CGC ATG CTC Asp Asn Ser Arg Met Leu 295	AAA GAC AAT TCG Lys Asp Asn Ser 300	1091
CTT TTC CAC ACT AAA GCC Leu Phe His Thr Lys Ala 305	ATG CCC ACT AAA AGC GTG Met Pro Thr Lys Ser Val 310	GAT GCG ATA ACT Asp Ala Ile Thr 315	1139
TCT CAA GCC AAA GAG CTT Ser Gln Ala Lys Glu Leu 320	AAC CAT TTA GTG GGG CAA Asn His Leu Val Gly Gln 325	ATC AAA GAA ATG Ile Lys Glu Met 330	1187
AAG CAA GAC GGG GCG AGT Lys Gln Asp Gly Ala Ser 340	CCT AGT AAG ATT GAT TCA Pro Ser Lys Ile Asp Ser 345	GTT GTC AAT AAA Val Val Asn Lys 350	1235
GCT ATG GAA GTG AGG GAC Ala Met Glu Val Arg Asp 355	AAG CTA GAC AAT AAT CTC Lys Leu Asp Asn Asn Leu 360	AAC CAA CTA GAC Asn Gln Leu Asp 365	1283
AAT GAC TTA AAA GAT CAA Asn Asp Leu Lys Asp Gln 370	AAA GGG CTT TCA AGC GAG Lys Lys Gly Leu Ser Ser 375	CAA CAA GCT CAA Gln Gln Gln Ala Gln 380	1331

-168-

370	375	380	
GTG GAT AAA GCC CTA GAC AGC Val Asp Lys Ala Leu Asp Ser 385	GTG CAA CAA TTA AGC CAT AGC AGC Val Gln Gln Leu Ser His Ser Ser 390	GTG 395	1379
GTG GTG GGG AAT TAT TTA GAC GGG AGT TTG AAA ATT GAT GGC GAT GAT Val Val Gly Asn Tyr Leu Asp Gly Ser Leu Lys Ile Asp Gly Asp Asp 400	405	410	415
AGA GAT GAT TTG AAT GAT GCG ATG AAT AAC CCT ATG CAA CAA CCC GTG Arg Asp Asp Leu Asn Asp Ala Met Asn Asn Pro Met Gln Gln Pro Val 420	425	430	1475
CAA CAA ACG CCT ACT AGC AAC ATG GCC GAC ACC CAT GCA AAT GAC AGC Gln Gln Thr Pro Thr Ser Asn Met Ala Asp Thr His Ala Asn Asp Ser 435	440	445	1523
AAG GAT CAA GGG AGT AAC GCG CTC ATA AAC CCT AAC AGC GCC ACT AAC Lys Asp Gln Gly Ser Asn Ala Leu Ile Asn Pro Asn Ser Ala Thr Asn 450	455	460	1571
GCC GAC GAC ACT CAC ACT GAC GAT ACT CAC ACT GAC ACT AAC ACC ACA Ala Asp Asp Thr His Thr Asp Asp Thr His Thr Asp Thr Asn Thr Thr 465	470	475	1619
AAC GAT GCT AGC ACC ACT GAC ACC CCC ACT GAC GAT AAA GAT GCT AGC Asn Asp Ala Ser Thr Thr Asp Thr Pro Thr Asp Asp Lys Asp Ala Ser 480	485	490	495
GGC TTG AAC AAT ACC GGC GAT ATG AAT AAC ACG GAT ACC GGC AAC ACG Gly Leu Asn Asn Thr Gly Asp Met Asn Asn Thr Asp Thr Gly Asn Thr 500	505	510	1715
GAC ACC GGC AAT ACG GAT ACC GGT AAC ACT GAT GAT ATG AGC AAC ATG Asp Thr Gly Asn Thr Asp Thr Gly Asn Thr Asp Asp Met Ser Asn Met 515	520	525	1763
AAC AAC GGC AAC GAT GAT ACG GGT AAC GCT AAT GAC GAC ATG AGC AAC Asn Asn Gly Asn Asp Asp Thr Gly Asn Ala Asn Asp Asp Met Ser Asn 530	535	540	1811
GGC AAC GAC ATG GGC GAT GAT TTG AAC AAC GCG AAC GAT ATG AAC GAC Gly Asn Asp Met Gly Asp Asp Leu Asn Asn Ala Asn Asp Met Asn Asp 545	550	555	1859
GAC ATG GGT AAT GGC AAC GAT GAC ATG GGC GAT ATG GGC GAT ATG AAC Asp Met Gly Asn Gly Asn Asp Asp Met Gly Asp Met Gly Asp Met Asn 560	565	570	575
GAC GAT ATG GGT GGC GAT ATG GGA GAC ATG GGG GAT ATG GGC GAT ATG Asp Asp Met Gly Gly Asp Met Gly Asp Met Gly Asp Met Gly Asp Met 580	585	590	1955
GGG AAT TGAGATTAAC CCCAATATCA AAGAGTGATA GCCAAACTT TAAGGAATAT TT			2013

-169-

Gly Asn

TTATAGTAAA AACGATTCTT TTAAGGTAAT AGGGGGGATA TTTTGC

2059

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 593 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Met	Lys	Ala	Ile	Lys	Ile	Leu	Phe	Ile	Met	Thr	Leu	Ser	Leu	Asn	Ala
1				5					10					15	
Ile	Ser	Val	Asn	Arg	Ala	Leu	Phe	Asp	Leu	Lys	Asp	Ser	Gln	Leu	Lys
			20					25					30		
Gly	Glu	Leu	Thr	Pro	Lys	Ile	Val	Asn	Phe	Gly	Gly	Tyr	Lys	Ser	Ser
		35				40						45			
Thr	Glu	Glu	Trp	Gly	Ala	Thr	Ala	Leu	Asn	Tyr	Ile	Asn	Ala	Ala	Asn
	50					55					60				
Gly	Asp	Ala	Lys	Lys	Phe	Ser	Thr	Leu	Val	Glu	Lys	Met	Arg	Phe	Asn
65				70					75					80	
Ser	Gly	Ile	Leu	Gly	Asn	Leu	Arg	Val	His	Ala	Arg	Leu	Arg	Gln	Ala
			85					90						95	
Leu	Lys	Leu	Gln	Lys	Asn	Leu	Lys	Tyr	Cys	Leu	Lys	Ile	Ile	Ala	Arg
		100						105					110		
Asp	Ser	Phe	Tyr	Ser	Tyr	Arg	Thr	Gly	Ile	Tyr	Ile	Pro	Leu	Gly	Ile
	115					120						125			
Ser	Leu	Lys	Asp	Gln	Lys	Thr	Ala	Gln	Lys	Met	Leu	Ala	Asp	Leu	Ser
	130					135					140				
Val	Val	Gly	Ala	Tyr	Leu	Lys	Lys	Gln	Gln	Glu	Asn	Glu	Lys	Ala	Gln
145				150				155						160	
Ser	Pro	Tyr	Tyr	Arg	Asn	Asn	Asn	Tyr	Tyr	Asn	Ser	Tyr	Tyr	Ser	Pro
			165					170						175	
Tyr	Tyr	Gly	Met	Tyr	Gly	Met	Tyr	Gly	Met	Gly	Met	Tyr	Gly	Met	Tyr
		180					185						190		
Gly	Met	Gly	Met	Tyr	Asp	Phe	Tyr	Asp	Phe	Tyr	Asp	Gly	Met	Tyr	Gly
	195					200						205			
Phe	Tyr	Pro	Asn	Met	Phe	Phe	Met	Met	Gln	Val	Gln	Asp	Tyr	Leu	Met
	210					215					220				
Leu	Glu	Asn	Tyr	Met	Tyr	Ala	Leu	Asp	Gln	Glu	Glu	Ile	Leu	Asp	His
225			230						235					240	
Asp	Ala	Ser	Thr	Asp	Gln	Leu	Asp	Thr	Pro	Thr	Asp	Asp	Asp	Lys	Asp
			245					250						255	
Asp	Lys	Asp	Asp	Lys	Ser	Leu	Gln	Gln	Ala	Asn	Leu	Met	Asn	Phe	Tyr
		260						265					270		
Arg	Asp	Pro	Lys	Phe	Ser	Lys	Gly	Ile	Gln	Thr	Asn	Arg	Leu	Asn	Ser
	275					280						285			
Ala	Leu	Val	Asn	Leu	Asp	Asn	Ser	Arg	Met	Leu	Lys	Asp	Asn	Ser	Leu

-170-

290		295		300
Phe His Thr Lys Ala Met	Pro Thr Lys Ser Val	Asp Ala Ile Thr Ser		
305	310	315	320	
Gln Ala Lys Glu Leu Asn	His Leu Val Gly Gln	Ile Lys Glu Met Lys		
	325	330	335	
Gln Asp Gly Ala Ser Pro	Ser Lys Ile Asp Ser	Val Val Asn Lys Ala		
	340	345	350	
Met Glu Val Arg Asp Lys	Leu Asp Asn Asn Leu	Asn Gln Leu Asp Asn		
	355	360	365	
Asp Leu Lys Asp Gln Lys	Gly Leu Ser Ser Glu	Gln Gln Ala Gln Val		
	370	375	380	
Asp Lys Ala Leu Asp Ser	Val Gln Gln Leu Ser	His Ser Ser Asp Val		
385	390	395	400	
Val Gly Asn Tyr Leu Asp	Gly Ser Leu Lys Ile	Asp Gly Asp Asp Arg		
	405	410	415	
Asp Asp Leu Asn Asp Ala	Met Asn Asn Pro Met	Gln Gln Pro Val Gln		
	420	425	430	
Gln Thr Pro Thr Ser Asn	Met Ala Asp Thr His	Ala Asn Asp Ser Lys		
	435	440	445	
Asp Gln Gly Ser Asn Ala	Leu Ile Asn Pro Asn	Ser Ala Thr Asn Ala		
	450	455	460	
Asp Asp Thr His Thr Asp	Asp Thr His Thr Asp	Thr Asn Thr Thr Asn		
465	470	475	480	
Asp Ala Ser Thr Thr Asp	Thr Pro Thr Asp Asp	Lys Asp Ala Ser Gly		
	485	490	495	
Leu Asn Asn Thr Gly Asp	Met Asn Asn Thr Asp	Thr Gly Asn Thr Asp		
	500	505	510	
Thr Gly Asn Thr Asp Thr	Gly Asn Thr Asp Asp	Met Ser Asn Met Asn		
	515	520	525	
Asn Gly Asn Asp Asp Thr	Gly Asn Ala Asn Asp	Asp Met Ser Asn Gly		
	530	535	540	
Asn Asp Met Gly Asp Asp	Leu Asn Asn Ala Asn	Asp Met Asn Asp Asp		
545	550	555	560	
Met Gly Asn Gly Asn Asp	Asp Met Gly Asp Met	Gly Asp Met Asn Asp		
	565	570	575	
Asp Met Gly Gly Asp Met	Gly Asp Met Gly Asp	Met Gly Asp Met Gly		
	580	585	590	
Asn				

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1527 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 112...1461
- (D) OTHER INFORMATION:

-171-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

AATGAGCGAT TTGAAAGATT TTGTCAATAA AACTTCAAGC CCTTTAAATG CGAATTGATT	60
TTCTTATATT ATGATTACGA TTTATCAATT TAAACATTT GGAGAAAGAC A ATG AGT	117
Met Ser	
1	
ATG GAA TTT GAT GCT GTT ATT ATT GGA GGT GGG GTT TCA GGG TGC GCG	165
Met Glu Phe Asp Ala Val Ile Ile Gly Gly Gly Val Ser Gly Cys Ala	
5 10 15	
ACC TTT TAT ACT TTG AGC GAA TAC AGC TCT TTA AAG CGC GTG GCT ATC	213
Thr Phe Tyr Thr Leu Ser Glu Tyr Ser Ser Leu Lys Arg Val Ala Ile	
20 25 30	
GTG GAA AAA TGC TCT AAA TTG GCT CAA ATC AGC TCC AGC GCT AAA GCT	261
Val Glu Lys Cys Ser Lys Leu Ala Gln Ile Ser Ser Ser Ala Lys Ala	
35 40 45 50	
AAT TCG CAA ACC ATT CAT GAT GGC TCT ATT GAA ACG AAT TAC ACT CCC	309
Asn Ser Gln Thr Ile His Asp Gly Ser Ile Glu Thr Asn Tyr Thr Pro	
55 60 65	
GAA AAA GCT AAA AAA GTG CGT TTG AGC GCT TAT AAG ACC AGG CAA TAC	357
Glu Lys Ala Lys Lys Val Arg Leu Ser Ala Tyr Lys Thr Arg Gln Tyr	
70 75 80	
GCT TTG AAT AAA GGC TTG CAA AAT GAA GTG ATT TTT GAA ACC CAG AAA	405
Ala Leu Asn Lys Gly Leu Gln Asn Glu Val Ile Phe Glu Thr Gln Lys	
85 90 95	
ATG GCT ATA GGC GTG GGC GAT GAA GAA TGC GAG TTC ATG AAA AAA CGC	453
Met Ala Ile Gly Val Gly Asp Glu Glu Cys Glu Phe Met Lys Lys Arg	
100 105 110	
TAC GAA TCT TTT AAA GAA ATC TTT GTG GGG TTA GAA GAA TTT GAC AAG	501
Tyr Glu Ser Phe Lys Glu Ile Phe Val Gly Leu Glu Glu Phe Asp Lys	
115 120 125 130	
CAA AAG ATT AAA GAA TTA GAG CCT AAT GTG ATT TTA GGG GCT AAT GGC	549
Gln Lys Ile Lys Glu Leu Glu Pro Asn Val Ile Leu Gly Ala Asn Gly	
135 140 145	
ATA GAC AGG CAT GAA AAC ATT ATC GGG CAT GGG TAT AGA AAG GAT TGG	597
Ile Asp Arg His Glu Asn Ile Ile Gly His Gly Tyr Arg Lys Asp Trp	
150 155 160	
AGC ACC ATG AAT TTT GCG AAG TTG AGT GAA AAC TTC GTT GAA GAA GCC	645
Ser Thr Met Asn Phe Ala Lys Leu Ser Glu Asn Phe Val Glu Glu Ala	
165 170 175	
CTA AAA TTA AAG CCT AAC AAC CAG GTG TTT TTG AAT TTC AAA GTG AAA	693
Leu Lys Leu Lys Pro Asn Asn Gln Val Phe Leu Asn Phe Lys Val Lys	
180 185 190	

-172-

AAG ATT GAA AAA CGC AAC GAC ACT TAC GCC GTA ATT TCA GAA GAC GCT Lys Ile Glu Lys Arg Asn Asp Thr Tyr Ala Val Ile Ser Glu Asp Ala 195 200 205 210	741
GAA GAA GTG TAT GCT AAA TTC GTG CTG GTC AAT GCC GGC TCT TAC GCT Glu Glu Val Tyr Ala Lys Phe Val Leu Val Asn Ala Gly Ser Tyr Ala 215 220 225	789
TTG CCT TTG GCT CAG AGC ATG GGC TAT GGC CTA GAT TTA GGG TGC TTG Leu Pro Leu Ala Gln Ser Met Gly Tyr Gly Leu Asp Leu Gly Cys Leu 230 235 240	837
CCT GTG GCG GGC AGC TTT TAT TTT GTG CCG GAT TTA TTA AGG GGT AAG Pro Val Ala Gly Ser Phe Tyr Phe Val Pro Asp Leu Leu Arg Gly Lys 245 250 255	885
GTT TAT ACC GTT CAA AAC CCC AAA CTC CCT TTT GCA GCC GTG CAT GGC Val Tyr Thr Val Gln Asn Pro Lys Leu Pro Phe Ala Ala Val His Gly 260 265 270	933
GAC CCT GAT GCC GTC ATT AAA GGA AAA ACA CGA ATC GGG CCT ACC GCT Asp Pro Asp Ala Val Ile Lys Gly Lys Thr Arg Ile Gly Pro Thr Ala 275 280 285 290	981
TTA ACG ATG CCT AAA TTA GAA CGC AAC AAA TGT TGG CTT AAG GGC ATT Leu Thr Met Pro Lys Leu Glu Arg Asn Lys Cys Trp Leu Lys Gly Ile 295 300 305	1029
AGC TTG GAA TTG TTG AAA ATG GAT TTG AAT AAA GAT GTG TTT AAA ATT Ser Leu Glu Leu Leu Lys Met Asp Leu Asn Lys Asp Val Phe Lys Ile 310 315 320	1077
GCG TTT GAT TTG ATG AGC GAT AAA GAA ATC CGA AAT TAT GTG TTT AAA Ala Phe Asp Leu Met Ser Asp Lys Glu Ile Arg Asn Tyr Val Phe Lys 325 330 335	1125
AAC ATG GTT TTT GAA TTG CCC ATT ATC GGT AAA AGG AAA TTT TTA AAA Asn Met Val Phe Glu Leu Pro Ile Ile Gly Lys Arg Lys Phe Leu Lys 340 345 350	1173
GAC GCT CAA AAA ATC ATC CCC TCT CTT AGC CTA GAA GAT CTA GAA TAC Asp Ala Gln Lys Ile Ile Pro Ser Leu Ser Leu Glu Asp Leu Glu Tyr 355 360 365 370	1221
GCT CAT GGT TTT GGT GAA GTG CGC CCG CAA GTT TTA GAC AGA ACC AAG Ala His Gly Phe Gly Glu Val Arg Pro Gln Val Leu Asp Arg Thr Lys 375 380 385	1269
CGA AAA CTG GAA TTA GGC GAA AAA AAG ATT TGC ACC CAT AAA GGC ATC Arg Lys Leu Glu Leu Gly Glu Lys Lys Ile Cys Thr His Lys Gly Ile 390 395 400	1317
ACT TTT AAC ATG ACC CCT TCT CCA GGC GCG ACG AGT TGT TTG CAA AAC Thr Phe Asn Met Thr Pro Ser Pro Gly Ala Thr Ser Cys Leu Gln Asn 405 410 415	1365

-173-

```

GCC CTT GTG GAT TCC CAA GAA ATC GCT GCG TAT TTG GGC GAG AGC TTT      1413
Ala Leu Val Asp Ser Gln Glu Ile Ala Ala Tyr Leu Gly Glu Ser Phe
   420               425               430

GAA TTA GAA CGC TTT TAT AAA GAT TTA TCC CCA GAA GAA TTG GAA AAT T      1462
Glu Leu Glu Arg Phe Tyr Lys Asp Leu Ser Pro Glu Glu Leu Glu Asn
   435               440               445               450

AAAAACGCAT GCAAAAAGAA CAAGAAGCCC AAGAAATCGC TAAAAAAGCC GTTAAAATCG      1522
TGTTT                                           1527

```

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```

Met Ser Met Glu Phe Asp Ala Val Ile Ile Gly Gly Gly Val Ser Gly
 1             5             10             15
Cys Ala Thr Phe Tyr Thr Leu Ser Glu Tyr Ser Ser Leu Lys Arg Val
   20             25             30
Ala Ile Val Glu Lys Cys Ser Lys Leu Ala Gln Ile Ser Ser Ser Ala
   35             40             45
Lys Ala Asn Ser Gln Thr Ile His Asp Gly Ser Ile Glu Thr Asn Tyr
   50             55             60
Thr Pro Glu Lys Ala Lys Lys Val Arg Leu Ser Ala Tyr Lys Thr Arg
   65             70             75             80
Gln Tyr Ala Leu Asn Lys Gly Leu Gln Asn Glu Val Ile Phe Glu Thr
   85             90             95
Gln Lys Met Ala Ile Gly Val Gly Asp Glu Glu Cys Glu Phe Met Lys
  100             105             110
Lys Arg Tyr Glu Ser Phe Lys Glu Ile Phe Val Gly Leu Glu Glu Phe
  115             120             125
Asp Lys Gln Lys Ile Lys Glu Leu Glu Pro Asn Val Ile Leu Gly Ala
  130             135             140
Asn Gly Ile Asp Arg His Glu Asn Ile Ile Gly His Gly Tyr Arg Lys
  145             150             155             160
Asp Trp Ser Thr Met Asn Phe Ala Lys Leu Ser Glu Asn Phe Val Glu
  165             170             175
Glu Ala Leu Lys Leu Lys Pro Asn Asn Gln Val Phe Leu Asn Phe Lys
  180             185             190
Val Lys Lys Ile Glu Lys Arg Asn Asp Thr Tyr Ala Val Ile Ser Glu
  195             200             205
Asp Ala Glu Glu Val Tyr Ala Lys Phe Val Leu Val Asn Ala Gly Ser
  210             215             220
Tyr Ala Leu Pro Leu Ala Gln Ser Met Gly Tyr Gly Leu Asp Leu Gly
  225             230             235             240
Cys Leu Pro Val Ala Gly Ser Phe Tyr Phe Val Pro Asp Leu Leu Arg
  245             250             255

```

-174-

Gly Lys Val Tyr Thr Val Gln Asn Pro Lys Leu Pro Phe Ala Ala Val
 260 265 270
 His Gly Asp Pro Asp Ala Val Ile Lys Gly Lys Thr Arg Ile Gly Pro
 275 280 285
 Thr Ala Leu Thr Met Pro Lys Leu Glu Arg Asn Lys Cys Trp Leu Lys
 290 295 300
 Gly Ile Ser Leu Glu Leu Leu Lys Met Asp Leu Asn Lys Asp Val Phe
 305 310 315 320
 Lys Ile Ala Phe Asp Leu Met Ser Asp Lys Glu Ile Arg Asn Tyr Val
 325 330 335
 Phe Lys Asn Met Val Phe Glu Leu Pro Ile Ile Gly Lys Arg Lys Phe
 340 345 350
 Leu Lys Asp Ala Gln Lys Ile Ile Pro Ser Leu Ser Leu Glu Asp Leu
 355 360 365
 Glu Tyr Ala His Gly Phe Gly Glu Val Arg Pro Gln Val Leu Asp Arg
 370 375 380
 Thr Lys Arg Lys Leu Glu Leu Gly Glu Lys Lys Ile Cys Thr His Lys
 385 390 395 400
 Gly Ile Thr Phe Asn Met Thr Pro Ser Pro Gly Ala Thr Ser Cys Leu
 405 410 415
 Gln Asn Ala Leu Val Asp Ser Gln Glu Ile Ala Ala Tyr Leu Gly Glu
 420 425 430
 Ser Phe Glu Leu Glu Arg Phe Tyr Lys Asp Leu Ser Pro Glu Glu Leu
 435 440 445
 Glu Asn
 450

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 63...590
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CTAGATTTAA TTTTAAAGTT ATATAATTAA ACCACAAAAT CCTTTTTTAA AAGAACTAA 60
 GC ATG CCA AAA CCC AAG AAA AAC ACC CTC CCC TGT AGC CTT TCT GTC 107
 Met Pro Lys Pro Lys Lys Asn Thr Leu Pro Cys Ser Leu Ser Val
 1 5 10 15
 AAA ATG TCT TAT TTC ATG CGC TTT CTC ATT AAA TGG CGC ACC CGC TCT 155
 Lys Met Ser Tyr Phe Met Arg Phe Leu Ile Lys Trp Arg Thr Arg Ser
 20 25 30
 TTA AGC CAT AAA ATG ATG ACT CTC ATT CAA ATC TTA AGC ATT CTG GCT 203

-175-

Leu	Ser	His	Lys	Met	Met	Thr	Leu	Ile	Gln	Ile	Leu	Ser	Ile	Leu	Ala		
			35				40					45					
TTA	GCG	AGC	AAG	GCC	AGT	GAA	GAT	TTA	GAA	GAG	CAA	CTC	AAA	AAA	ATC	251	
Leu	Ala	Ser	Lys	Ala	Ser	Glu	Asp	Leu	Glu	Glu	Gln	Leu	Lys	Lys	Ile		
			50				55					60					
AAA	GAT	TAC	ATT	TAT	AGA	ACC	CTA	AAC	GCT	AAA	ATC	GCA	TCG	GAT	GTG	299	
Lys	Asp	Tyr	Ile	Tyr	Arg	Thr	Leu	Asn	Ala	Lys	Ile	Ala	Ser	Asp	Val		
	65					70					75						
TAT	AAC	CGA	GTG	CTT	ATT	TTA	GTG	AAT	GAA	TAT	TGC	ACT	AAT	GAA	GAA	347	
Tyr	Asn	Arg	Val	Leu	Ile	Leu	Val	Asn	Glu	Tyr	Cys	Thr	Asn	Glu	Glu		
80					85				90					95			
TTG	TTT	GAC	AAA	GAG	AGC	GTT	AAA	ATT	TCA	GAT	TTA	CTC	ATT	CAA	GAC	395	
Leu	Phe	Asp	Lys	Glu	Ser	Val	Lys	Ile	Ser	Asp	Leu	Leu	Ile	Gln	Asp		
				100					105					110			
ATT	CAG	CTT	TAC	GCT	TTA	GTG	GAT	GAA	ATG	CTT	AAA	GAA	GAT	AAA	TAT	443	
Ile	Gln	Leu	Tyr	Ala	Leu	Val	Asp	Glu	Met	Leu	Lys	Glu	Asp	Lys	Tyr		
			115					120					125				
CAA	GTC	CAG	CAC	ACC	ATT	TTA	AAG	GGC	ATC	ATC	AAA	CGC	AAA	TAC	GAT	491	
Gln	Val	Gln	His	Thr	Ile	Leu	Lys	Gly	Ile	Ile	Lys	Arg	Lys	Tyr	Asp		
			130					135					140				
GAA	GCC	TAC	TCG	CTC	AAT	AGC	GAA	GAC	AGG	ATT	CTT	TTA	GAA	TAC	CAA	539	
Glu	Ala	Tyr	Ser	Leu	Asn	Ser	Glu	Asp	Arg	Ile	Leu	Leu	Glu	Tyr	Gln		
			145			150					155						
GAA	CGC	TTG	CTA	GAA	CAC	TCA	CAC	GCG	TCT	TTT	TCA	AAT	AAA	AAA	TTC	587	
Glu	Arg	Leu	Leu	Glu	His	Ser	His	Ala	Ser	Phe	Ser	Asn	Lys	Lys	Phe		
160					165				170					175			
AAA	TGATTTGAAA	GCGTTACTTG	CCCTGCTTTT	TGGGCTTTTA	TTGAAAAAGG	GCTTTA	646										
Lys																	
AAATGAG																653	

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Met	Pro	Lys	Pro	Lys	Lys	Asn	Thr	Leu	Pro	Cys	Ser	Leu	Ser	Val	Lys
1				5					10					15	

-176-

```

Met Ser Tyr Phe Met Arg Phe Leu Ile Lys Trp Arg Thr Arg Ser Leu
      20                      25                      30
Ser His Lys Met Met Thr Leu Ile Gln Ile Leu Ser Ile Leu Ala Leu
      35                      40                      45
Ala Ser Lys Ala Ser Glu Asp Leu Glu Glu Gln Leu Lys Lys Ile Lys
      50                      55                      60
Asp Tyr Ile Tyr Arg Thr Leu Asn Ala Lys Ile Ala Ser Asp Val Tyr
      65                      70                      75                      80
Asn Arg Val Leu Ile Leu Val Asn Glu Tyr Cys Thr Asn Glu Glu Leu
      85                      90                      95
Phe Asp Lys Glu Ser Val Lys Ile Ser Asp Leu Leu Ile Gln Asp Ile
      100                     105                     110
Gln Leu Tyr Ala Leu Val Asp Glu Met Leu Lys Glu Asp Lys Tyr Gln
      115                     120                     125
Val Gln His Thr Ile Leu Lys Gly Ile Ile Lys Arg Lys Tyr Asp Glu
      130                     135                     140
Ala Tyr Ser Leu Asn Ser Glu Asp Arg Ile Leu Leu Glu Tyr Gln Glu
      145                     150                     155                     160
Arg Leu Leu Glu His Ser His Ala Ser Phe Ser Asn Lys Lys Phe Lys
      165                     170                     175

```

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1883 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 91...1833
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```

AAGCGTTAAA TTCCAATCAA AAACCATCGT ATCGGTGTTA ATATTGTGTA AAAATTAATG      60
TTATGAATCT CTTGTATTAA AAGGACTTCA ATG AAA AAA TTG GTT TTA GTC ATC      114
Met Lys Lys Leu Val Leu Val Ile
      1                      5

TTT TTA ACG CTA GCG CTT TCA ATA TCT GCA AAA GAA GTC AAA ATA GTG      162
Phe Leu Thr Leu Ala Leu Ser Ile Ser Ala Lys Glu Val Lys Ile Val
      10                      15                      20

TTT TTA GAA ACT TCA GAC ATT CAT GGG CGG CTT TTT TCG TAT GAT TAT      210
Phe Leu Glu Thr Ser Asp Ile His Gly Arg Leu Phe Ser Tyr Asp Tyr
      25                      30                      35                      40

GCG ATT GGC GAG CAA AAA CCC AAT AAC GGC TTG ACA AGG ATT GCG ACT      258
Ala Ile Gly Glu Gln Lys Pro Asn Asn Gly Leu Thr Arg Ile Ala Thr
      45                      50                      55

```

-177-

TTA ATC AAA AAG CAA AGG GCT GAG AAT AAA AAT GTG GTT TTG ATT GAC	306
Leu Ile Lys Lys Gln Arg Ala Glu Asn Lys Asn Val Val Leu Ile Asp	
60 65 70	
AGC GGG GAT TTG TTG CAA GGC AAT AGC GCG GAG TTG TTT AAT GAT GAG	354
Ser Gly Asp Leu Leu Gln Gly Asn Ser Ala Glu Leu Phe Asn Asp Glu	
75 80 85	
CCA ATT CAT CCG CTA GTT AGA GCT GAA AAC GAT TTG AAA TTT GAC ATT	402
Pro Ile His Pro Leu Val Arg Ala Glu Asn Asp Leu Lys Phe Asp Ile	
90 95 100	
CGT GTG CTT GGC AAT CAC GAG TTT AAT TTC AGT AAA GAT TTT TTA GAA	450
Arg Val Leu Gly Asn His Glu Phe Asn Phe Ser Lys Asp Phe Leu Glu	
105 110 115 120	
AAG AAT ATT AAG GGG TTT AAT GGC GAT GTC ATG AAT GCG AAT ATC ATT	498
Lys Asn Ile Lys Gly Phe Asn Gly Asp Val Met Asn Ala Asn Ile Ile	
125 130 135	
AAA ATT GCG GAC AAT AAG CCG TTT GTA AAA CCT TAT ATT ATT AAA AAA	546
Lys Ile Ala Asp Asn Lys Pro Phe Val Lys Pro Tyr Ile Ile Lys Lys	
140 145 150	
ATT GAT GGC GTG AGG GTG GCG GTT GTG GGG TAT GTG GTG GCG CAC ATC	594
Ile Asp Gly Val Arg Val Ala Val Val Gly Tyr Val Val Ala His Ile	
155 160 165	
CCA ACT TGG GAG GCC TCT ACG CCT GAA CAT TTT GCA GGA TTG AAG TTT	642
Pro Thr Trp Glu Ala Ser Thr Pro Glu His Phe Ala Gly Leu Lys Phe	
170 175 180	
TTG GAC GCT GAA GAA GCG TTA AAA AAG ACC TTA AAA GAG TTG AAA GGG	690
Leu Asp Ala Glu Glu Ala Leu Lys Lys Thr Leu Lys Glu Leu Lys Gly	
185 190 195 200	
AAG TAT GAT ATT TTG ATT GGC GCT TTT CAT TTG GGG CGA GAA GAT GAG	738
Lys Tyr Asp Ile Leu Ile Gly Ala Phe His Leu Gly Arg Glu Asp Glu	
205 210 215	
AAA GGT GGC GAC GGG ATA CCG GAT TTA GCG AAA AAA TTC CCG CAA TTT	786
Lys Gly Gly Asp Gly Ile Pro Asp Leu Ala Lys Lys Phe Pro Gln Phe	
220 225 230	
GAC ATC ATT TTT GCA GGG CAT GAG CAT GCG GTT TAT AAC ACC AAA GTA	834
Asp Ile Ile Phe Ala Gly His Glu His Ala Val Tyr Asn Thr Lys Val	
235 240 245	
GGG AAA GTG CAT ACC ATT GAG CCT GGA GCG TAT GGG GCT TAT CTG GCA	882
Gly Lys Val His Thr Ile Glu Pro Gly Ala Tyr Gly Ala Tyr Leu Ala	
250 255 260	
AAG GGC GTG GTG GTA TTT GAC ACT AAA ACG AAG AAA AAA ATT ATA ACG	930
Lys Gly Val Val Val Phe Asp Thr Lys Thr Lys Lys Lys Ile Ile Thr	
265 270 275 280	

-178-

ACT GAA AAT TTA CCC ACA AAA GAT GTG CCA GAA GAT GAA GAA TTA GCG	978
Thr Glu Asn Leu Pro Thr Lys Asp Val Pro Glu Asp Glu Glu Leu Ala	
285 290 295	
AAA AAA TAC GAA TAT GTG GAT AAA AAA TCA AAA GAA TAC GCT AAT GAA	1026
Lys Lys Tyr Glu Tyr Val Asp Lys Lys Ser Lys Glu Tyr Ala Asn Glu	
300 305 310	
GTG GTT GGC GAA GTT ACA AAA ACC TTT ATT GAC AGG CCT GAT TTT ATC	1074
Val Val Gly Glu Val Thr Lys Thr Phe Ile Asp Arg Pro Asp Phe Ile	
315 320 325	
ACA GGA GAA GAA AAA ATC ACC ACG ATG CCC ACC GCC GCC TTG CAA GAA	1122
Thr Gly Glu Glu Lys Ile Thr Thr Met Pro Thr Ala Ala Leu Gln Glu	
330 335 340	
ACA CCG GTG ATA GAA TTG ATT AAT AAA GTG CAA AAA TAT TAC GCA AAA	1170
Thr Pro Val Ile Glu Leu Ile Asn Lys Val Gln Lys Tyr Tyr Ala Lys	
345 350 355 360	
GCC GAT GTT TCA GCG GCA GCC TTA TTC AAT TTT GGG GCG AAT TTG AAA	1218
Ala Asp Val Ser Ala Ala Ala Leu Phe Asn Phe Gly Ala Asn Leu Lys	
365 370 375	
AAA GGG CCT TTC AAA AGA AAA GAT GTC ACT TAT ATT TAC AAG TTC GCT	1266
Lys Gly Pro Phe Lys Arg Lys Asp Val Thr Tyr Ile Tyr Lys Phe Ala	
380 385 390	
AAT ACG CTC ATT GGA GTG CGT ATA ACG GGT GAA AAT CTG TTG AAA TAC	1314
Asn Thr Leu Ile Gly Val Arg Ile Thr Gly Glu Asn Leu Leu Lys Tyr	
395 400 405	
ATG GAA TGG TCA TAC CGA TTT TAC AAT CAG TTG CAA CCA GGA GAT TTG	1362
Met Glu Trp Ser Tyr Arg Phe Tyr Asn Gln Leu Gln Pro Gly Asp Leu	
410 415 420	
ACG ATC AGT TTT AAT GAA AAC ATT CGC GGC TAT AAC TTT GAT ATG TTT	1410
Thr Ile Ser Phe Asn Glu Asn Ile Arg Gly Tyr Asn Phe Asp Met Phe	
425 430 435 440	
TCT GGC GTG AAA TAC CAG GTT GAT GTT ACA AAA CCC GCC GGA CAA AGG	1458
Ser Gly Val Lys Tyr Gln Val Asp Val Thr Lys Pro Ala Gly Gln Arg	
445 450 455	
ATT ATC AAT CCG ACA ATC AAC AAC AAA CCC ATT GAC CCC AAA GCC ATC	1506
Ile Ile Asn Pro Thr Ile Asn Asn Lys Pro Ile Asp Pro Lys Ala Ile	
460 465 470	
TAT AAA TTA GCG ATC AAC AAT TAC CGA TTC GGA ACA TTA TCC ACG ACA	1554
Tyr Lys Leu Ala Ile Asn Asn Tyr Arg Phe Gly Thr Leu Ser Thr Thr	
475 480 485	
TTG AAT TTG GTT ACA GAC GCT GMT AGG TAT TAT AAT TCT TAC GAT GAA	1602
Leu Asn Leu Val Thr Asp Ala Xaa Arg Tyr Tyr Asn Ser Tyr Asp Glu	
490 495 500	

CTG	CAA	GAT	AAT	GGG	CAA	ATA	CGA	GAT	TTG	ATC	ATC	AAA	TAC	ATC	ACG	1650
Leu	Gln	Asp	Asn	Gly	Gln	Ile	Arg	Asp	Leu	Ile	Ile	Lys	Tyr	Ile	Thr	
				510				515				520				
GAA	GAA	AAA	GGT	GGG	AAG	GTA	ACC	CCT	GAA	TTG	GAG	GGT	AAT	TGG	GAA	1698
Glu	Glu	Lys	Gly	Gly	Lys	Val	Thr	Pro	Glu	Leu	Glu	Gly	Asn	Trp	Glu	
				525				530				535				
ATC	ATC	AAC	TAC	GAT	TTC	AAA	AAC	CCG	TTG	TTG	GAA	AAA	TTG	AGA	GAA	1746
Ile	Ile	Asn	Tyr	Asp	Phe	Lys	Asn	Pro	Leu	Leu	Glu	Lys	Leu	Arg	Glu	
				540				545				550				
AAA	TTA	AAA	GAG	GGG	AGC	ATC	AAA	ATC	CCC	ACC	TCA	AAG	GAT	GGG	AGG	1794
Lys	Leu	Lys	Glu	Gly	Ser	Ile	Lys	Ile	Pro	Thr	Ser	Lys	Asp	Gly	Arg	
				555				560				565				
ACT	TTG	AAT	GTC	AAA	TCC	ATT	AAA	GAG	AGT	GAA	GTT	AAA	TAA AATT			1840
Thr	Leu	Asn	Val	Lys	Ser	Ile	Lys	Glu	Ser	Glu	Val	Lys				
				570				575				580				
TTTTATTTTTT ATTATTTTAT CTTTAAGCCT AAC TTA AAAAA AGG												1883				

(A) LENGTH: 581 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(v) FRAGMENT TYPE: internal

Met	Lys	Lys	Leu	Val	Leu	Val	Ile	Phe	Leu	Thr	Leu	Ala	Leu	Ser	Ile
1				5					10					15	
Ser	Ala	Lys	Glu	Val	Lys	Ile	Val	Phe	Leu	Glu	Thr	Ser	Asp	Ile	His
			20					25					30		
Gly	Arg	Leu	Phe	Ser	Tyr	Asp	Tyr	Ala	Ile	Gly	Glu	Gln	Lys	Pro	Asn
		35					40					45			
Asn	Gly	Leu	Thr	Arg	Ile	Ala	Thr	Leu	Ile	Lys	Lys	Gln	Arg	Ala	Glu
	50					55					60				
Asn	Lys	Asn	Val	Val	Leu	Ile	Asp	Ser	Gly	Asp	Leu	Leu	Gln	Gly	Asn
65					70					75					80
Ser	Ala	Glu	Leu	Phe	Asn	Asp	Glu	Pro	Ile	His	Pro	Leu	Val	Arg	Ala
				85					90					95	
Glu	Asn	Asp	Leu	Lys	Phe	Asp	Ile	Arg	Val	Leu	Gly	Asn	His	Glu	Phe
			100					105					110		
Asn	Phe	Ser	Lys	Asp	Phe	Leu	Glu	Lys	Asn	Ile	Lys	Gly	Phe	Asn	Gly
		115					120					125			
Asp	Val	Met	Asn	Ala	Asn	Ile	Ile	Lys	Ile	Ala	Asp	Asn	Lys	Pro	Phe
	130					135					140				
Val	Lys	Pro	Tyr	Ile	Ile	Lys	Lys	Ile	Asp	Gly	Val	Arg	Val	Ala	Val

-180-

145					150					155				160
Val	Gly	Tyr	Val	Val	Ala	His	Ile	Pro	Thr	Trp	Glu	Ala	Ser	Thr
				165					170					175
Glu	His	Phe	Ala	Gly	Leu	Lys	Phe	Leu	Asp	Ala	Glu	Glu	Ala	Leu
			180					185					190	
Lys	Thr	Leu	Lys	Glu	Leu	Lys	Gly	Lys	Tyr	Asp	Ile	Leu	Ile	Gly
			195				200					205		
Phe	His	Leu	Gly	Arg	Glu	Asp	Glu	Lys	Gly	Gly	Asp	Gly	Ile	Pro
	210					215				220				
Leu	Ala	Lys	Lys	Phe	Pro	Gln	Phe	Asp	Ile	Ile	Phe	Ala	Gly	His
225					230				235					240
His	Ala	Val	Tyr	Asn	Thr	Lys	Val	Gly	Lys	Val	His	Thr	Ile	Glu
				245					250					255
Gly	Ala	Tyr	Gly	Ala	Tyr	Leu	Ala	Lys	Gly	Val	Val	Val	Phe	Asp
			260					265					270	
Lys	Thr	Lys	Lys	Lys	Ile	Ile	Thr	Thr	Glu	Asn	Leu	Pro	Thr	Lys
	275						280					285		
Val	Pro	Glu	Asp	Glu	Glu	Leu	Ala	Lys	Lys	Tyr	Glu	Tyr	Val	Asp
	290					295					300			
Lys	Ser	Lys	Glu	Tyr	Ala	Asn	Glu	Val	Val	Gly	Glu	Val	Thr	Lys
305					310					315				320
Phe	Ile	Asp	Arg	Pro	Asp	Phe	Ile	Thr	Gly	Glu	Glu	Lys	Ile	Thr
				325					330					335
Met	Pro	Thr	Ala	Ala	Leu	Gln	Glu	Thr	Pro	Val	Ile	Glu	Leu	Ile
			340					345					350	
Lys	Val	Gln	Lys	Tyr	Tyr	Ala	Lys	Ala	Asp	Val	Ser	Ala	Ala	Ala
	355					360						365		
Phe	Asn	Phe	Gly	Ala	Asn	Leu	Lys	Lys	Gly	Pro	Phe	Lys	Arg	Lys
	370					375					380			
Val	Thr	Tyr	Ile	Tyr	Lys	Phe	Ala	Asn	Thr	Leu	Ile	Gly	Val	Arg
385					390					395				400
Thr	Gly	Glu	Asn	Leu	Leu	Lys	Tyr	Met	Glu	Trp	Ser	Tyr	Arg	Phe
			405						410					415
Asn	Gln	Leu	Gln	Pro	Gly	Asp	Leu	Thr	Ile	Ser	Phe	Asn	Glu	Asn
			420					425					430	
Arg	Gly	Tyr	Asn	Phe	Asp	Met	Phe	Ser	Gly	Val	Lys	Tyr	Gln	Val
	435					440						445		
Val	Thr	Lys	Pro	Ala	Gly	Gln	Arg	Ile	Ile	Asn	Pro	Thr	Ile	Asn
	450					455					460			
Lys	Pro	Ile	Asp	Pro	Lys	Ala	Ile	Tyr	Lys	Leu	Ala	Ile	Asn	Asn
465					470					475				480
Arg	Phe	Gly	Thr	Leu	Ser	Thr	Thr	Leu	Asn	Leu	Val	Thr	Asp	Ala
			485						490					495
Arg	Tyr	Tyr	Asn	Ser	Tyr	Asp	Glu	Leu	Gln	Asp	Asn	Gly	Gln	Ile
			500					505					510	
Asp	Leu	Ile	Ile	Lys	Tyr	Ile	Thr	Glu	Glu	Lys	Gly	Gly	Lys	Val
	515						520					525		
Pro	Glu	Leu	Glu	Gly	Asn	Trp	Glu	Ile	Ile	Asn	Tyr	Asp	Phe	Lys
	530					535					540			
Pro	Leu	Leu	Glu	Lys	Leu	Arg	Glu	Lys	Leu	Lys	Glu	Gly	Ser	Ile
545					550					555				560
Ile	Pro	Thr	Ser	Lys	Asp	Gly	Arg	Thr	Leu	Asn	Val	Lys	Ser	Ile
				565					570					575
Glu	Ser	Glu	Val	Lys										
			580											

-181-

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1339 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 68...1252
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```

CCAATCGTTT AATAGCGATT AAATATGACT ATATACACTA CAACAATAAG ATTTTGAAAG      60
GTTGGTA ATG GAA TCA GTA AAA ACA GGA AAA ACA AAT AAG GTT GGC AAG      109
    Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys
        1             5             10

AAT ACA GAG ATG GCT AAT ACA AAG GCA AAT AAA GAG GCT CAT TTT AAA      157
Asn Thr Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys
15             20             25             30

CAA GCG AGC ACC ATT ACA AAT ATA ATC AGA TCA ATT CGT GGG ATT TTT      205
Gln Ala Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe
        35             40             45

ACA AAA ATT GCA AAG AAA GTT AGA GGA CTT GTA AAA AAA CAC CCC AAG      253
Thr Lys Ile Ala Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys
        50             55             60

AAA AGC AGT GCG GCA TTA GTA GTA TTG ACC CAT ATT GCG TGC AAG AAA      301
Lys Ser Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys
        65             70             75

GCG AAA GAA TTA GAC GAT AAA GTC CAA GAT AAA TCC AAA CAA GCT GAA      349
Ala Lys Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu
        80             85             90

AAA GAA AAT CAA ATC AAT TGG TGG AAA TAT TCA GGA TTA ACA ATA GCG      397
Lys Glu Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala
        95             100            105            110

ACA AGT TTA TTA TTA GCC GCT TGT AGC ACT GGT GAT GTT AGT GAA CAA      445
Thr Ser Leu Leu Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln
        115            120            125

ATA GAA CTA GAA CAA GAA AAA CAA AAG ACG AGC AAT ATA GAG ACT AAC      493
Ile Glu Leu Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn
        130            135            140

```

-182-

AAT CAA ATA AAA GTA GAA CAA GAA AAA CAA AAG ACA AGC AAT ATA GAG	541
Asn Gln Ile Lys Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu	
145 150 155	
ACT AAT AAT CAA ATA AAA GTA GAA CAA GAA CAA CAA AAG ACA AGC AAT	589
Thr Asn Asn Gln Ile Lys Val Glu Gln Glu Gln Gln Lys Thr Ser Asn	
160 165 170	
ACA CAG AAA GAT TTG GTT AAA GAA CAG AAA GAT TTG GTT AAA GAA CAG	637
Thr Gln Lys Asp Leu Val Lys Glu Gln Lys Asp Leu Val Lys Glu Gln	
175 180 185 190	
AAA GAT TTG GTT AAA GAA CAG AAA GAT TTG GTT AAA GAA CAG AAA GAT	685
Lys Asp Leu Val Lys Glu Gln Lys Asp Leu Val Lys Glu Gln Lys Asp	
195 200 205	
TTG GTT AAA ACA CAG AAA GAT TTC ATT AAA TAT GTA GAA CAA AAT TGC	733
Leu Val Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys	
210 215 220	
CAA GAA AAT CAT AAT CAA TTC TTT ATT GAA AAA GGA GGA ATT AAG GCT	781
Gln Glu Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala	
225 230 235	
GGT ATT GGT ATA GAA GTA GAA GCT GAA TGC AAA ACC CCT AAA CCT GCA	829
Gly Ile Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala	
240 245 250	
AAA ACC AAT CAA ACC CCT ATC CAG CCA AAA CAC CTC CCA AAC TCT AAA	877
Lys Thr Asn Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys	
255 260 265 270	
CAA CCC CGC TCT CAA AGA GGA TCA AAA GCG CAA GAG CTT ATC GCT TAT	925
Gln Pro Arg Ser Gln Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr	
275 280 285	
TTG CAA AAA GAG CTA GAA TTT CTG CCC TAT TCG CAA AAA GCT ATC GCT	973
Leu Gln Lys Glu Leu Glu Phe Leu Pro Tyr Ser Gln Lys Ala Ile Ala	
290 295 300	
AAA CAA GTG GAT TTT TAC AGG CCA AGT TCT ATC GCT TAT TTA GAA CTA	1021
Lys Gln Val Asp Phe Tyr Arg Pro Ser Ser Ile Ala Tyr Leu Glu Leu	
305 310 315	
GAT CCT AGA GAT TTT AAG GTT ACA GAA GAA TGG CAA AAA GAA AAT CTA	1069
Asp Pro Arg Asp Phe Lys Val Thr Glu Glu Trp Gln Lys Glu Asn Leu	
320 325 330	
AAA ATA CGC TCT AAA GCT CAA GCT AAA ATG CTT GAA ATG AGA AAC CCA	1117
Lys Ile Arg Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Asn Pro	
335 340 345 350	
CAA GCC CAC CTT TCA AAC TCT CAA AGC CTT TTG TTC GTT CAA AAA ATA	1165
Gln Ala His Leu Ser Asn Ser Gln Ser Leu Leu Phe Val Gln Lys Ile	
355 360 365	

-183-

```

TTT GCT GAT GTT AAT AAA GAA ATA GAA GCA GTT GCT AAT ACT GAA AAG      1213
Phe Ala Asp Val Asn Lys Glu Ile Glu Ala Val Ala Asn Thr Glu Lys
      370                      375                      380

AAA GCA GAA AAA GCG GGT TAT GGT TAT AGT AAA AGG ATG TAGCGGTTAA AA    1264
Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met
      385                      390                      395

ACATTGCACC AAGTTTTTTAA TTATCTGTCG GCTTTTGAAA ACATTTTTTTA TGGTAGCGTT    1324
ATTTGGCAAT AAAAG                                                         1339

```

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 395 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

```

Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr
 1          5          10          15
Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Ala His Phe Lys Gln Ala
 20          25          30
Ser Thr Ile Thr Asn Ile Ile Arg Ser Ile Arg Gly Ile Phe Thr Lys
 35          40          45
Ile Ala Lys Lys Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser
 50          55          60
Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys
 65          70          75          80
Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu
 85          90          95
Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser
100          105          110
Leu Leu Leu Ala Ala Cys Ser Thr Gly Asp Val Ser Glu Gln Ile Glu
115          120          125
Leu Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln
130          135          140
Ile Lys Val Glu Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn
145          150          155          160
Asn Gln Ile Lys Val Glu Gln Glu Gln Gln Lys Thr Ser Asn Thr Gln
165          170          175
Lys Asp Leu Val Lys Glu Gln Lys Asp Leu Val Lys Glu Gln Lys Asp
180          185          190
Leu Val Lys Glu Gln Lys Asp Leu Val Lys Glu Gln Lys Asp Leu Val
195          200          205
Lys Thr Gln Lys Asp Phe Ile Lys Tyr Val Glu Gln Asn Cys Gln Glu
210          215          220
Asn His Asn Gln Phe Phe Ile Glu Lys Gly Gly Ile Lys Ala Gly Ile
225          230          235          240
Gly Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr

```


					245					250					255
Asn	Gln	Thr	Pro	Ile	Gln	Pro	Lys	His	Leu	Pro	Asn	Ser	Lys	Gln	Pro
			260					265					270		
Arg	Ser	Gln	Arg	Gly	Ser	Lys	Ala	Gln	Glu	Leu	Ile	Ala	Tyr	Leu	Gln
		275					280					285			
Lys	Glu	Leu	Glu	Phe	Leu	Pro	Tyr	Ser	Gln	Lys	Ala	Ile	Ala	Lys	Gln
	290					295					300				
Val	Asp	Phe	Tyr	Arg	Pro	Ser	Ser	Ile	Ala	Tyr	Leu	Glu	Leu	Asp	Pro
305					310					315					320
Arg	Asp	Phe	Lys	Val	Thr	Glu	Glu	Trp	Gln	Lys	Glu	Asn	Leu	Lys	Ile
			325						330					335	
Arg	Ser	Lys	Ala	Gln	Ala	Lys	Met	Leu	Glu	Met	Arg	Asn	Pro	Gln	Ala
			340					345					350		
His	Leu	Ser	Asn	Ser	Gln	Ser	Leu	Leu	Phe	Val	Gln	Lys	Ile	Phe	Ala
		355					360					365			
Asp	Val	Asn	Lys	Glu	Ile	Glu	Ala	Val	Ala	Asn	Thr	Glu	Lys	Lys	Ala
	370					375					380				
Glu	Lys	Ala	Gly	Tyr	Gly	Tyr	Ser	Lys	Arg	Met					
385					390					395					

(A) LENGTH: 904 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 70...864
(D) OTHER INFORMATION:

BNSDOCID: <WO__9821225A1_I_>

-185-

TGT ATA GCG TTA AAA ATT GGA GCG AGC GAT AAA GTA GCG AGC GCT AAT	303
Cys Ile Ala Leu Lys Ile Gly Ala Ser Asp Lys Val Ala Ser Ala Asn	
65 70 75	
AAG GCT ATT TCG TGG GGT TTG AGC GGG TAT GCG AAT AAA ACG GTG TGG	351
Lys Ala Ile Ser Trp Gly Leu Ser Gly Tyr Ala Asn Lys Thr Val Trp	
80 85 90	
CTT GTG TTT GTC GCT TTC ATT TTG GGT TTA GGG TAT GAA AAA AGC TTG	399
Leu Val Phe Val Ala Phe Ile Leu Gly Leu Gly Tyr Glu Lys Ser Leu	
95 100 105 110	
TTA GGG AAA CGG ATC GCT CTT TTA CTG ATT AGG TTT TTA GGG CAA ACC	447
Leu Gly Lys Arg Ile Ala Leu Leu Ile Arg Phe Leu Gly Gln Thr	
115 120 125	
CCT TTA GGT TTA GGC TAT GCG ATT GGT TTG AGC GAA TTG TGT CTA GCC	495
Pro Leu Gly Leu Gly Tyr Ala Ile Gly Leu Ser Glu Leu Cys Leu Ala	
130 135 140	
CCT TTT ATC CCT AGC AAC TCC GCT AGA AGT GGA GGC ATA CTC TAT CCC	543
Pro Phe Ile Pro Ser Asn Ser Ala Arg Ser Gly Gly Ile Leu Tyr Pro	
145 150 155	
ATC GTT TCA TCT ATC CCG CCT TTA ATG GGA TCT ACT CCA AAT AAT AAC	591
Ile Val Ser Ser Ile Pro Pro Leu Met Gly Ser Thr Pro Asn Asn Asn	
160 165 170	
CCT GAC AAA ATC GGC GCG TAT TTG ATG TGG GTC GCT TTG GCT TCA ACT	639
Pro Asp Lys Ile Gly Ala Tyr Leu Met Trp Val Ala Leu Ala Ser Thr	
175 180 185 190	
TGC ATC ACT TCG TCC ATG TTT TTA ACC GCG CTC GCT CCT AAC CCC CTA	687
Cys Ile Thr Ser Ser Met Phe Leu Thr Ala Leu Ala Pro Asn Pro Leu	
195 200 205	
GCA ATG GAA ATC GCT GCC AAA ATG GGC GTG AAT GAA ATC TCA TGG TTT	735
Ala Met Glu Ile Ala Ala Lys Met Gly Val Asn Glu Ile Ser Trp Phe	
210 215 220	
TCG TGG TTT TTA GCG TTC TTG CCT TGT GGG GTG GTT TTG ATC TTG CTT	783
Ser Trp Phe Leu Ala Phe Leu Pro Cys Gly Val Val Leu Ile Leu Leu	
225 230 235	
GTG CCT TTA TTG GCG TAT AAA ACC TGC AAA CCC ACC TTA AAA GGC TCA	831
Val Pro Leu Leu Ala Tyr Lys Thr Cys Lys Pro Thr Leu Lys Gly Ser	
240 245 250	
AAA GAA GTG AGT TTG TGG GCC AAA AAA AGG AAT TAGAGGGCAT GGGGAGGTTT	884
Lys Glu Val Ser Leu Trp Ala Lys Lys Arg Asn	
255 260 265	
TCTTTAAAAG AAATTTTAAT	904

(2) INFORMATION FOR SEQ ID NO:76:

-186-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

Met Ile Lys Gln Thr Leu Ile Ile Leu Ala Pro Phe Phe Ile Ala Thr
 1           5           10           15
Leu Leu Tyr Phe Leu Gly Ala Pro Asp Gly Leu Arg Pro Asn Ala Trp
           20           25           30
Leu Tyr Phe Cys Ile Phe Met Gly Met Ile Ile Gly Leu Ile Leu Glu
           35           40           45
Pro Val Pro Ser Gly Leu Ile Ala Leu Ser Ala Leu Val Leu Cys Ile
           50           55           60
Ala Leu Lys Ile Gly Ala Ser Asp Lys Val Ala Ser Ala Asn Lys Ala
65           70           75           80
Ile Ser Trp Gly Leu Ser Gly Tyr Ala Asn Lys Thr Val Trp Leu Val
           85           90           95
Phe Val Ala Phe Ile Leu Gly Leu Gly Tyr Glu Lys Ser Leu Leu Gly
           100          105          110
Lys Arg Ile Ala Leu Leu Leu Ile Arg Phe Leu Gly Gln Thr Pro Leu
           115          120          125
Gly Leu Gly Tyr Ala Ile Gly Leu Ser Glu Leu Cys Leu Ala Pro Phe
130          135          140
Ile Pro Ser Asn Ser Ala Arg Ser Gly Gly Ile Leu Tyr Pro Ile Val
145          150          155          160
Ser Ser Ile Pro Pro Leu Met Gly Ser Thr Pro Asn Asn Asn Pro Asp
           165          170          175
Lys Ile Gly Ala Tyr Leu Met Trp Val Ala Leu Ala Ser Thr Cys Ile
           180          185          190
Thr Ser Ser Met Phe Leu Thr Ala Leu Ala Pro Asn Pro Leu Ala Met
           195          200          205
Glu Ile Ala Ala Lys Met Gly Val Asn Glu Ile Ser Trp Phe Ser Trp
210          215          220
Phe Leu Ala Phe Leu Pro Cys Gly Val Val Leu Ile Leu Leu Val Pro
225          230          235          240
Leu Leu Ala Tyr Lys Thr Cys Lys Pro Thr Leu Lys Gly Ser Lys Glu
           245          250          255
Val Ser Leu Trp Ala Lys Lys Arg Asn
           260          265

```

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

-187-

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 152...1069
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TTTTAAGCGG	TTTCCCTAAA	ATAGGTTTTT	AATCAATTTA	ATCCAAAGTT	GAATTTATTT	60
TTTGACAATA	TTATACTATA	ATAACCAATT	AGATTGGGGT	TTTACTGATT	TTTCTTTGTG	120
TGAGCTTTGG	CTTAGTTTTG	TAAGGAATGA	G ATG ATA AAG AGT TGG ACT AAA			172
			Met Ile Lys Ser Trp Thr Lys			
			1	5		
AAG TGG TTT TTG ATT TTA TTT TTA ATG GCA AGT TGT TCC AGT TAT TTG						220
Lys Trp Phe Leu Ile Leu Phe Leu Met Ala Ser Cys Ser Ser Tyr Leu						
	10		15		20	
GTG GCT ACA ACC GGT GAG AAA TAT TTT AAA ATG GCT ACT CAA GCC TTT						268
Val Ala Thr Thr Gly Glu Lys Tyr Phe Lys Met Ala Thr Gln Ala Phe						
	25		30		35	
AAG AGA GGG GAC TAC CAT AAA GCG GTG GCT TTT TAT AAG AGG AGC TGT						316
Lys Arg Gly Asp Tyr His Lys Ala Val Ala Phe Tyr Lys Arg Ser Cys						
	40		45		50	55
AAT TTA AGG GTG GGG GTT GGT TGC ACG AGT TTA GGC TCT ATG TAT GAA						364
Asn Leu Arg Val Gly Val Gly Cys Thr Ser Leu Gly Ser Met Tyr Glu						
		60		65		70
GAT GGC GAT GGC GTG GAT CAG AAT ATT ACA AAA GCC GTT TTT TAT TAC						412
Asp Gly Asp Gly Val Asp Gln Asn Ile Thr Lys Ala Val Phe Tyr Tyr						
	75		80		85	
AGA AGA GGG TGT AAT TTA AGG AAT CAT CTC GCT TGC GCG AGT CTA GGC						460
Arg Arg Gly Cys Asn Leu Arg Asn His Leu Ala Cys Ala Ser Leu Gly						
	90		95		100	
TCT ATG TAT GAA GAT GGC GAT GGT GTG CAA AAA AAC CTT CCA AAG GCT						508
Ser Met Tyr Glu Asp Gly Asp Gly Val Gln Lys Asn Leu Pro Lys Ala						
	105		110		115	
ATC TAT TAT TAC AGG AGA GGG TGC CAC TTA AAG GGT GGG GTG AGC TGT						556
Ile Tyr Tyr Tyr Arg Arg Gly Cys His Leu Lys Gly Gly Val Ser Cys						
	120		125		130	135
GGG AGT TTA GGT TTT ATG TAT TTT AAT GGC ACG GGC GTT AAG CAA AAT						604
Gly Ser Leu Gly Phe Met Tyr Phe Asn Gly Thr Gly Val Lys Gln Asn						
	140		145		150	
TAT GCC AAA GCC CTT TTT CTT TCT AAA TAC GCT TGC AGT TTG AAT TAC						652
Tyr Ala Lys Ala Leu Phe Leu Ser Lys Tyr Ala Cys Ser Leu Asn Tyr						
	155		160		165	
GGC ATT AGT TGT AAC TTT GTA GGG TAT ATG TAT AGG AAC GCC AAA GGC						700

-188-

Gly Ile Ser Cys Asn Phe Val Gly Tyr Met Tyr Arg Asn Ala Lys Gly	
170 175 180	
GTA CAG AAG GAT TTG AAA AAA GCC CTT GCG AAT TTT AAA AGA GGG TGC	748
Val Gln Lys Asp Leu Lys Lys Ala Leu Ala Asn Phe Lys Arg Gly Cys	
185 190 195	
CAT TTG AAA GAC GGA GCG AGT TGT GTG AGC TTG GGA TAC ATG TAT GAA	796
His Leu Lys Asp Gly Ala Ser Cys Val Ser Leu Gly Tyr Met Tyr Glu	
200 205 210 215	
GTC GGT ATG GAT GTC AAA CAA AAT GGA GAG CAA GCC TTG AAT CTT TAT	844
Val Gly Met Asp Val Lys Gln Asn Gly Glu Gln Ala Leu Asn Leu Tyr	
220 225 230	
AAA AAG GGT TGT TAT TTA AAA AGG GGG AGC GGT TGT CAT AAT GTG GCG	892
Lys Lys Gly Cys Tyr Leu Lys Arg Gly Ser Gly Cys His Asn Val Ala	
235 240 245	
GTG ATG TAT TAC ACC GGT AAG GGC GTT CCA AAG GAT TTA GAT AAA GCC	940
Val Met Tyr Tyr Thr Gly Lys Gly Val Pro Lys Asp Leu Asp Lys Ala	
250 255 260	
ATT TCG TAT TAT AAG AAA GGT TGC ACT CTA GGC TTT AGT GGT AGC TGT	988
Ile Ser Tyr Tyr Lys Lys Gly Cys Thr Leu Gly Phe Ser Gly Ser Cys	
265 270 275	
AAA GTG TTA GAA GAA GTG ATT GGC AAG AAG TCT GAT GAT TTG CAA GAT	1036
Lys Val Leu Glu Glu Val Ile Gly Lys Lys Ser Asp Asp Leu Gln Asp	
280 285 290 295	
GAC GCG CAA AAC GAC ACG CAA GAT GAT ATG CAA TAAGTTAAAG CTTATGGACT	1089
Asp Ala Gln Asn Asp Thr Gln Asp Asp Met Gln	
300 305	
AATGATTAAA ACTCATCTTA TAGAAATCTT TCTACTCTCT TGTTATCAAA TAGGGATTAA	1149
GCGTCTCTAT TGATGGGTAT TGAGACTAAA AATCTGCAAA TCTAG	1194

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Met Ile Lys Ser Trp Thr Lys Lys Trp Phe Leu Ile Leu Phe Leu Met
1 5 10 15
Ala Ser Cys Ser Ser Tyr Leu Val Ala Thr Thr Gly Glu Lys Tyr Phe
20 25 30
Lys Met Ala Thr Gln Ala Phe Lys Arg Gly Asp Tyr His Lys Ala Val

- 189 -

[illegible]

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1001 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 101...865
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

TTTGTTTATA AGAAAAATTA TTTCAAATGT AGTAGAATTA AGGCAGTGTT TTTGCGTCAA 60
GCGATTTTAT GTTAATTTTG AGTTTTTATG AGCAGTTTTT ATG CAA CAA GAA GAG 115

-190-

															Met	Gln	Gln	Glu	Glu	
															1				5	
ATT	ATA	GAG	GGT	TAT	TAT	GGT	GCT	AGC	AAA	GGG	CTT	AAA	AAG	AGC	GGT	163				
Ile	Ile	Glu	Gly	Tyr	Tyr	Gly	Ala	Ser	Lys	Gly	Leu	Lys	Lys	Ser	Gly					
				10						15						20				
ATT	TAT	GCC	AAG	CTG	GAT	TTT	TTA	CAG	AGC	GCT	ACG	GGC	TTG	ATT	TTA	211				
Ile	Tyr	Ala	Lys	Leu	Asp	Phe	Leu	Gln	Ser	Ala	Thr	Gly	Leu	Ile	Leu					
			25						30						35					
GCG	CTC	TTT	ATG	ATA	GCA	CAC	ATG	TTT	TTA	GTC	TCA	AGT	ATC	TTG	ATT	259				
Ala	Leu	Phe	Met	Ile	Ala	His	Met	Phe	Leu	Val	Ser	Ser	Ile	Leu	Ile					
			40						45						50					
AGC	GAT	GAA	GCC	ATG	TAT	AAA	GTG	GCG	AAA	TTT	TTT	GAA	GGG	AGC	TTG	307				
Ser	Asp	Glu	Ala	Met	Tyr	Lys	Val	Ala	Lys	Phe	Phe	Glu	Gly	Ser	Leu					
		55						60						65						
TTT	TTA	AAA	GCG	GGC	GAG	CCG	GCT	ATT	GTG	AGC	GTG	GTT	GCA	GCA	GGG	355				
Phe	Leu	Lys	Ala	Gly	Glu	Pro	Ala	Ile	Val	Ser	Val	Val	Ala	Ala	Gly					
70						75						80		85						
ATT	ATT	CTT	ATT	TTA	GTC	GCG	CAT	GCT	TTT	TTG	GCG	TTA	AGG	AAA	TTC	403				
Ile	Ile	Leu	Ile	Leu	Val	Ala	His	Ala	Phe	Leu	Ala	Leu	Arg	Lys	Phe					
				90						95						100				
CCT	ATC	AAT	TAC	AGG	CAA	TAC	AAG	GTT	TTT	AAA	ACC	CAT	AAG	CAT	TTG	451				
Pro	Ile	Asn	Tyr	Arg	Gln	Tyr	Lys	Val	Phe	Lys	Thr	His	Lys	His	Leu					
			105						110						115					
ATG	AAA	CAT	GGC	GAT	ACG	AGC	TTG	TGG	TTT	ATT	CAA	GCC	CTC	ACC	GGG	499				
Met	Lys	His	Gly	Asp	Thr	Ser	Leu	Trp	Phe	Ile	Gln	Ala	Leu	Thr	Gly					
			120						125						130					
TTT	GCG	ATG	TTT	TTC	TTA	GCG	AGT	ATC	CAC	TTA	TTT	GTC	ATG	CTC	ACA	547				
Phe	Ala	Met	Phe	Phe	Leu	Ala	Ser	Ile	His	Leu	Phe	Val	Met	Leu	Thr					
135						140						145								
GAG	CCT	GAA	AGT	ATT	GGG	CCT	CAT	GGT	TCA	AGC	TAT	CGT	TTT	GTA	ACG	595				
Glu	Pro	Glu	Ser	Ile	Gly	Pro	His	Gly	Ser	Ser	Tyr	Arg	Phe	Val	Thr					
150						155						160		165						
CAA	AAC	TTT	TGG	CTT	TTG	TAT	ATT	TTC	TTA	TTG	TTT	GCC	GTA	GAA	TTG	643				
Gln	Asn	Phe	Trp	Leu	Leu	Tyr	Ile	Phe	Leu	Leu	Phe	Ala	Val	Glu	Leu					
				170						175						180				
CAT	GGC	TCT	ATT	GGG	TTG	TAT	CGT	TTA	GCG	ATC	AAA	TGG	GGG	TGG	TTT	691				
His	Gly	Ser	Ile	Gly	Leu	Tyr	Arg	Leu	Ala	Ile	Lys	Trp	Gly	Trp	Phe					
			185						190						195					
AAA	AAT	GTG	AGC	ATT	CAA	GGT	TTG	AGA	AAA	GTC	AAA	TGG	GCG	ATG	AGC	739				
Lys	Asn	Val	Ser	Ile	Gln	Gly	Leu	Arg	Lys	Val	Lys	Trp	Ala	Met	Ser					
			200						205						210					
GTG	TTT	TTT	ATT	GTT	TTA	GGG	CTT	TGC	ACC	TAT	GGG	GCT	TAC	ATT	AAA	787				

-191-

Val Phe Phe Ile Val Leu Gly Leu Cys Thr Tyr Gly Ala Tyr Ile Lys
 215 220 225

AAA GGT TTA GAA AAT AAG GAA AAT GGC ATT AAA ACC ATG CAA GAA GCC 835
 Lys Gly Leu Glu Asn Lys Glu Asn Gly Ile Lys Thr Met Gln Glu Ala
 230 235 240 245

ATA GAA GCT GAT GGG AAA TTC CAC AAA GAA TAAGGGTAGA AAATGAAAAT AAC 888
 Ile Glu Ala Asp Gly Lys Phe His Lys Glu
 250 255

ATATTGTGAT GCGCTAATTA TTGGAGGCGG ACTAGCTGGG TTAAGGGCTA GTATCGCATG 948
 CAAACAAAAG GGTTTAAACA CCATCGTTTT AAGCCTAGTG CCTGTCAGGC GTT 1001

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 255 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Met Gln Gln Glu Glu Ile Ile Glu Gly Tyr Tyr Gly Ala Ser Lys Gly
 1 5 10 15

Leu Lys Lys Ser Gly Ile Tyr Ala Lys Leu Asp Phe Leu Gln Ser Ala
 20 25 30

Thr Gly Leu Ile Leu Ala Leu Phe Met Ile Ala His Met Phe Leu Val
 35 40 45

Ser Ser Ile Leu Ile Ser Asp Glu Ala Met Tyr Lys Val Ala Lys Phe
 50 55 60

Phe Glu Gly Ser Leu Phe Leu Lys Ala Gly Glu Pro Ala Ile Val Ser
 65 70 75 80

Val Val Ala Ala Gly Ile Ile Leu Ile Leu Val Ala His Ala Phe Leu
 85 90 95

Ala Leu Arg Lys Phe Pro Ile Asn Tyr Arg Gln Tyr Lys Val Phe Lys
 100 105 110

Thr His Lys His Leu Met Lys His Gly Asp Thr Ser Leu Trp Phe Ile
 115 120 125

Gln Ala Leu Thr Gly Phe Ala Met Phe Phe Leu Ala Ser Ile His Leu
 130 135 140

Phe Val Met Leu Thr Glu Pro Glu Ser Ile Gly Pro His Gly Ser Ser
 145 150 155 160

Tyr Arg Phe Val Thr Gln Asn Phe Trp Leu Leu Tyr Ile Phe Leu Leu
 165 170 175

Phe Ala Val Glu Leu His Gly Ser Ile Gly Leu Tyr Arg Leu Ala Ile
 180 185 190

Lys Trp Gly Trp Phe Lys Asn Val Ser Ile Gln Gly Leu Arg Lys Val
 195 200 205

Lys Trp Ala Met Ser Val Phe Phe Ile Val Leu Gly Leu Cys Thr Tyr
 210 215 220

-192-

Gly	Ala	Tyr	Ile	Lys	Lys	Gly	Leu	Glu	Asn	Lys	Glu	Asn	Gly	Ile	Lys
225					230					235					240
Thr	Met	Gln	Glu	Ala	Ile	Glu	Ala	Asp	Gly	Lys	Phe	His	Lys	Glu	
				245					250					255	

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 975 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 82...912
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

TTTTAAAATT AAAGAAAATT TTTTAAAG ATTATCACTC TTTTGTGATA AAGTAATCAT	60
TTAAAATTTA GGGAGTTTTT T ATG GAA GAA TCA ACA GCG TTT ATT TTG GCT	111
Met Glu Glu Ser Thr Ala Phe Ile Leu Ala	
1 5 10	
CTT GTG GGG CTA TTC ACC GGC ATT ACC GCC GGG TTT TTT GGT ATT GGT	159
Leu Val Gly Leu Phe Thr Gly Ile Thr Ala Gly Phe Phe Gly Ile Gly	
15 20 25	
GGG GGG GAG ATT GTC GTC CCT AGC GCG ATT TTT GCC CAT TTT AGC TAT	207
Gly Gly Glu Ile Val Val Pro Ser Ala Ile Phe Ala His Phe Ser Tyr	
30 35 40	
AGC CAT GCG GTG GGT ATT TCG CTC ATG CAA ATG CTT TTT TCT TCA GTG	255
Ser His Ala Val Gly Ile Ser Leu Met Gln Met Leu Phe Ser Ser Val	
45 50 55	
GTC GGC TCT ATC ATC AAT TAC AAA AAG GGC TTA TTG GAT TTG AGA GAA	303
Val Gly Ser Ile Ile Asn Tyr Lys Lys Gly Leu Leu Asp Leu Arg Glu	
60 65 70	
GGC TCA TTT GCC GCG CTT GGA GGG CTA ATG GGA GCG ATT TTA GGG AGC	351
Gly Ser Phe Ala Ala Leu Gly Gly Leu Met Gly Ala Ile Leu Gly Ser	
75 80 85 90	
TTT ATC TTA AAA ATC ATT GAC GAT AAA ATT TTA ATG GCG GTG TTT GTG	399
Phe Ile Leu Lys Ile Ile Asp Asp Lys Ile Leu Met Ala Val Phe Val	
95 100 105	
GTG GTG GTG TGC TAC ACC TTT ATC AAA TAC GCT TTT TCT AGC AAC AAG	447
Val Val Val Cys Tyr Thr Phe Ile Lys Tyr Ala Phe Ser Ser Asn Lys	
110 115 120	

-193-

AAA CCC AAG CAT TTT GAA GAA ATG CAT TTT GAT TTG CAT GCG AAT AAC	495
Lys Pro Lys His Phe Glu Glu Met His Phe Asp Leu His Ala Asn Asn	
125 130 135	
AAA ACG CCC GAA AAA AAG CGC GCA ATC CCT TTT GTG TCT ATG GAT AGA	543
Lys Thr Pro Glu Lys Lys Arg Ala Ile Pro Phe Val Ser Met Asp Arg	
140 145 150	
ACG CAT GGG GTT TTG ATG CTC GCC GGT TTT GTT ACC GGC ATC TTT TCT	591
Thr His Gly Val Leu Met Leu Ala Gly Phe Val Thr Gly Ile Phe Ser	
155 160 165 170	
ATC CCA CTA GGC ATG GGT GGA GGG ATT TTA ATG GTG CCG TTT TTG GGC	639
Ile Pro Leu Gly Met Gly Gly Gly Ile Leu Met Val Pro Phe Leu Gly	
175 180 185	
TAT TTT TTG AAA TAC GAT TCT AAA AAA ATC GTG CCT TTG GGG CTA TTT	687
Tyr Phe Leu Lys Tyr Asp Ser Lys Lys Ile Val Pro Leu Gly Leu Phe	
190 195 200	
TTT GTG GTG TTC GCT TCT TTA TCT GGG GTC ATC TCT CTT TAT AAC GGG	735
Phe Val Val Phe Ala Ser Leu Ser Gly Val Ile Ser Leu Tyr Asn Gly	
205 210 215	
AGG GTT CTT GAT AAT ATA AGC GTT CAA GCG GGG GTG ATT ACC GGC ATT	783
Arg Val Leu Asp Asn Ile Ser Val Gln Ala Gly Val Ile Thr Gly Ile	
220 225 230	
GGA GCG TTT TTA GGC GTG GGC ATT GGC ATC AAG CTT ATC GCT TTG GCT	831
Gly Ala Phe Leu Gly Val Gly Ile Gly Ile Lys Leu Ile Ala Leu Ala	
235 240 245 250	
AAT GAA AAG GTG CAT AAA ATC CTG TTG CTC CTT ATT TAT GCT TTA AGC	879
Asn Glu Lys Val His Lys Ile Leu Leu Leu Leu Ile Tyr Ala Leu Ser	
255 260 265	
ATT TTA GCG ACT TTA CAC AAG CTC ATT ATG GGG TAAATCTAAA AACGCTTCTA	932
Ile Leu Ala Thr Leu His Lys Leu Ile Met Gly	
270 275	
GGGCATTTTT AAAATTAATA TCAAAGAGCT TTCACCAGCA AGC	975

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

-194-

```

Met Glu Glu Ser Thr Ala Phe Ile Leu Ala Leu Val Gly Leu Phe Thr
 1           5           10           15
Gly Ile Thr Ala Gly Phe Phe Gly Ile Gly Gly Gly Glu Ile Val Val
          20           25           30
Pro Ser Ala Ile Phe Ala His Phe Ser Tyr Ser His Ala Val Gly Ile
          35           40           45
Ser Leu Met Gln Met Leu Phe Ser Ser Val Val Gly Ser Ile Ile Asn
          50           55           60
Tyr Lys Lys Gly Leu Leu Asp Leu Arg Glu Gly Ser Phe Ala Ala Leu
65           70           75           80
Gly Gly Leu Met Gly Ala Ile Leu Gly Ser Phe Ile Leu Lys Ile Ile
          85           90           95
Asp Asp Lys Ile Leu Met Ala Val Phe Val Val Val Val Cys Tyr Thr
          100           105           110
Phe Ile Lys Tyr Ala Phe Ser Ser Asn Lys Lys Pro Lys His Phe Glu
          115           120           125
Glu Met His Phe Asp Leu His Ala Asn Asn Lys Thr Pro Glu Lys Lys
          130           135           140
Arg Ala Ile Pro Phe Val Ser Met Asp Arg Thr His Gly Val Leu Met
145           150           155           160
Leu Ala Gly Phe Val Thr Gly Ile Phe Ser Ile Pro Leu Gly Met Gly
          165           170           175
Gly Gly Ile Leu Met Val Pro Phe Leu Gly Tyr Phe Leu Lys Tyr Asp
          180           185           190
Ser Lys Lys Ile Val Pro Leu Gly Leu Phe Phe Val Val Phe Ala Ser
          195           200           205
Leu Ser Gly Val Ile Ser Leu Tyr Asn Gly Arg Val Leu Asp Asn Ile
          210           215           220
Ser Val Gln Ala Gly Val Ile Thr Gly Ile Gly Ala Phe Leu Gly Val
225           230           235           240
Gly Ile Gly Ile Lys Leu Ile Ala Leu Ala Asn Glu Lys Val His Lys
          245           250           255
Ile Leu Leu Leu Leu Ile Tyr Ala Leu Ser Ile Leu Ala Thr Leu His
          260           265           270
Lys Leu Ile Met Gly
          275

```

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1667 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 220...1482
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

-195-

AAGCGCGAGC TATATGAGGA ATTTTAGCTT CTATGTGGGC TATTCAGTCG GTTTTTAAGG	60
AAGGCTCTTG ATGAAAAATA CCAATACAAA AGAGATAAAG AATACAAGAA TGAAAAAAGG	120
TTATAGTCAA TACCACGCGC TCAAAAAAGG GCTTTTAAAA ACGCTCTGCT TTTTAGCCTT	180
CCTTTAAGCG TGGCGTTAGC TGAAGACGAT GGCTTTTAT ATG GGA GTG GGC TAT	234
Met Gly Val Gly Tyr	
1 5	
CAA ATC GGC GGC GCG CAA CAA AAT ATC GAT AAC AAA GGC AGC ACC CTA	282
Gln Ile Gly Gly Ala Gln Gln Asn Ile Asp Asn Lys Gly Ser Thr Leu	
10 15 20	
AGG AAT AAT GTC ATT AAT AAT TTC CGC CAA GTG GGC GTG GGT ATG GCA	330
Arg Asn Asn Val Ile Asn Asn Phe Arg Gln Val Gly Val Gly Met Ala	
25 30 35	
GGG GGT AAT GGG CTT TTA GCC TTA GCG ACA AAC ACG ACC ATG GAC GCT	378
Gly Gly Asn Gly Leu Leu Ala Leu Ala Thr Asn Thr Thr Met Asp Ala	
40 45 50	
CTT TTA GGG ATA GGC AAC CAA ATT GTC AAT ACT AAT ACA ACT GTT AGC	426
Leu Leu Gly Ile Gly Asn Gln Ile Val Asn Thr Asn Thr Thr Val Ser	
55 60 65	
AAC AAC AAC GCA GAA TTA ACC CAG TTT AAA AAA ATA CTC CCT CAA ATT	474
Asn Asn Asn Ala Glu Thr Gln Phe Lys Lys Ile Leu Pro Gln Ile	
70 75 80 85	
GAG CAA CGC TTT GAA ACG AAT AAA AAC GCT TAT AGC GTT CAA GCC TTG	522
Glu Gln Arg Phe Glu Thr Asn Lys Asn Ala Tyr Ser Val Gln Ala Leu	
90 95 100	
CAA GTG TAT TTG AGT AAT GTG CTT TAT AAC TTG GTT AAT AAT AGT AAT	570
Gln Val Tyr Leu Ser Asn Val Leu Tyr Asn Leu Val Asn Asn Ser Asn	
105 110 115	
AAT GGC AGT AAT AAT GGA GTC GTT CCT GAA TAT GTA GGA ATT ATA AAA	618
Asn Gly Ser Asn Asn Gly Val Val Pro Glu Tyr Val Gly Ile Ile Lys	
120 125 130	
GTT CTC TAT GGT TCT CAA AAT GAA TTC AGT CTC TTA GCC ACG GAG AGT	666
Val Leu Tyr Gly Ser Gln Asn Glu Phe Ser Leu Leu Ala Thr Glu Ser	
135 140 145	
GTG GTG CTT TTA AAC GCG CTT ACA AGG GTG AAT CTG GAT AGT AAT TCG	714
Val Val Leu Leu Asn Ala Leu Thr Arg Val Asn Leu Asp Ser Asn Ser	
150 155 160 165	
GTG TTT TTA AAA GGG CTA TTA GCC CAA ATG CAG CTT TTT AAT GAC ACT	762
Val Phe Leu Lys Gly Leu Leu Ala Gln Met Gln Leu Phe Asn Asp Thr	
170 175 180	
TCT TCA GCA AAG CTA GGC CAG ATC GCA GAA AAC TTG AAG AAC GGT GGT	810
Ser Ser Ala Lys Leu Gly Gln Ile Ala Glu Asn Leu Lys Asn Gly Gly	
185 190 195	
GCA GGA TCA ATG CTC CAA AAG GAT GTG AAA ACC ATC TCG GAT CGA ATC	858

-196-

Ala Gly Ser Met Leu Gln Lys Asp Val Lys Thr Ile Ser Asp Arg Ile	
200 205 210	
GCT ACT TAC CAA GAG AAT CTA AAA CAG CTA GGA GGG ATG CTA AAG AAT	906
Ala Thr Tyr Gln Glu Asn Leu Lys Gln Leu Gly Gly Met Leu Lys Asn	
215 220 225	
TAC GAT GAA CCC TAC TTG CCC CAA TTT GGG CCA GGC ACA AGC TCT CAG	954
Tyr Asp Glu Pro Tyr Leu Pro Gln Phe Gly Pro Gly Thr Ser Ser Gln	
230 235 240 245	
CAT GGG GTT ATT AAT GGC TTT GGC ATT CAA GTG GGC TAT AAG CAA TTT	1002
His Gly Val Ile Asn Gly Phe Gly Ile Gln Val Gly Tyr Lys Gln Phe	
250 255 260	
TTT GGG AAC AAG CGG AAT ATA GGC TTA CGA TAT TAC GCT TTC TTT GAT	1050
Phe Gly Asn Lys Arg Asn Ile Gly Leu Arg Tyr Tyr Ala Phe Phe Asp	
265 270 275	
TAT GGC TTT ACG CAA TTG GGC AGT CTT AGC AGC GCC GTT AAA GCG AAT	1098
Tyr Gly Phe Thr Gln Leu Gly Ser Leu Ser Ser Ala Val Lys Ala Asn	
280 285 290	
ATC TTT ACT TAT GGC GCT GGC ACG GAC TTT TTA TGG AAT ATC TTT AGA	1146
Ile Phe Thr Tyr Gly Ala Gly Thr Asp Phe Leu Trp Asn Ile Phe Arg	
295 300 305	
AGG GTT TTT AGC GAT CAG TCC TTG AAT GTG GGG GTG TTT GGG GGC ATT	1194
Arg Val Phe Ser Asp Gln Ser Leu Asn Val Gly Val Phe Gly Gly Ile	
310 315 320 325	
CAA ATA GCG GGT AAC ACT TGG GAT AGC TCT TTA AGA GGT CAA ATT GAA	1242
Gln Ile Ala Gly Asn Thr Trp Asp Ser Ser Leu Arg Gly Gln Ile Glu	
330 335 340	
AAC TCG TTT AAA GAA TAC CCC ACT CCC ACG AAT TTC CAA TTT TTG TTT	1290
Asn Ser Phe Lys Glu Tyr Pro Thr Pro Thr Asn Phe Gln Phe Leu Phe	
345 350 355	
AAT TTG GGT TTA AGG GCT CAT TTT GCC AGC ACC ATG CAC CGC CGG TTT	1338
Asn Leu Gly Leu Arg Ala His Phe Ala Ser Thr Met His Arg Arg Phe	
360 365 370	
TTG AGC GCG TCT CAA AGC ATT CAG CAT GGG ATG GAA TTT GGC GTG AAA	1386
Leu Ser Ala Ser Gln Ser Ile Gln His Gly Met Glu Phe Gly Val Lys	
375 380 385	
ATC CCG GCT ATC AAT CAA AGG TAT TTG AGG GCC AAT GGG GCT GAT GTG	1434
Ile Pro Ala Ile Asn Gln Arg Tyr Leu Arg Ala Asn Gly Ala Asp Val	
390 395 400 405	
GAT TAC AGG CGT TTG TAT GCG TTC TAT ATC AAT TAC ACG ATA GGT TTT T	1483
Asp Tyr Arg Arg Leu Tyr Ala Phe Tyr Ile Asn Tyr Thr Ile Gly Phe	
410 415 420	

-197-

```

AAGCTCTTTT TAGGGCTTAT AAAGAGGCTT TTTACTTTTT TTTTGGTATT CTAACAAGCT 1543
TTTAAATAAT CCAATCTACT TTGTTTTAAG GATAATATTT TATGGCAGAT GTCGTTGTGG 1603
GGATCCAGTG GGGAGATGAG GGAAGGGAA AAATTGTTGA TAGGATCGCT AAAGATTATG 1663
ACTT 1667

```

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 421 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

```

Met Gly Val Gly Tyr Gln Ile Gly Gly Ala Gln Gln Asn Ile Asp Asn
 1           5           10           15
Lys Gly Ser Thr Leu Arg Asn Asn Val Ile Asn Asn Phe Arg Gln Val
          20           25           30
Gly Val Gly Met Ala Gly Gly Asn Gly Leu Leu Ala Leu Ala Thr Asn
          35           40           45
Thr Thr Met Asp Ala Leu Leu Gly Ile Gly Asn Gln Ile Val Asn Thr
          50           55           60
Asn Thr Thr Val Ser Asn Asn Ala Glu Leu Thr Gln Phe Lys Lys
65           70           75           80
Ile Leu Pro Gln Ile Glu Gln Arg Phe Glu Thr Asn Lys Asn Ala Tyr
          85           90           95
Ser Val Gln Ala Leu Gln Val Tyr Leu Ser Asn Val Leu Tyr Asn Leu
          100          105          110
Val Asn Asn Ser Asn Asn Gly Ser Asn Asn Gly Val Val Pro Glu Tyr
          115          120          125
Val Gly Ile Ile Lys Val Leu Tyr Gly Ser Gln Asn Glu Phe Ser Leu
          130          135          140
Leu Ala Thr Glu Ser Val Val Leu Leu Asn Ala Leu Thr Arg Val Asn
145          150          155          160
Leu Asp Ser Asn Ser Val Phe Leu Lys Gly Leu Leu Ala Gln Met Gln
          165          170          175
Leu Phe Asn Asp Thr Ser Ser Ala Lys Leu Gly Gln Ile Ala Glu Asn
          180          185          190
Leu Lys Asn Gly Gly Ala Gly Ser Met Leu Gln Lys Asp Val Lys Thr
          195          200          205
Ile Ser Asp Arg Ile Ala Thr Tyr Gln Glu Asn Leu Lys Gln Leu Gly
          210          215          220
Gly Met Leu Lys Asn Tyr Asp Glu Pro Tyr Leu Pro Gln Phe Gly Pro
225          230          235          240
Gly Thr Ser Ser Gln His Gly Val Ile Asn Gly Phe Gly Ile Gln Val
          245          250          255
Gly Tyr Lys Gln Phe Phe Gly Asn Lys Arg Asn Ile Gly Leu Arg Tyr
          260          265          270
Tyr Ala Phe Phe Asp Tyr Gly Phe Thr Gln Leu Gly Ser Leu Ser Ser
          275          280          285
Ala Val Lys Ala Asn Ile Phe Thr Tyr Gly Ala Gly Thr Asp Phe Leu

```

-198-

290		295		300	
Trp Asn Ile Phe Arg Arg Val Phe Ser Asp Gln Ser Leu Asn Val Gly					
305		310		315	320
Val Phe Gly Gly Ile Gln Ile Ala Gly Asn Thr Trp Asp Ser Ser Leu					
	325		330		335
Arg Gly Gln Ile Glu Asn Ser Phe Lys Glu Tyr Pro Thr Pro Thr Asn					
	340		345		350
Phe Gln Phe Leu Phe Asn Leu Gly Leu Arg Ala His Phe Ala Ser Thr					
	355		360		365
Met His Arg Arg Phe Leu Ser Ala Ser Gln Ser Ile Gln His Gly Met					
	370		375		380
Glu Phe Gly Val Lys Ile Pro Ala Ile Asn Gln Arg Tyr Leu Arg Ala					
385		390		395	400
Asn Gly Ala Asp Val Asp Tyr Arg Arg Leu Tyr Ala Phe Tyr Ile Asn					
	405		410		415
Tyr Thr Ile Gly Phe					
	420				

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 926 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 207...746
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

CCCCCTTAATT GCAGATGTTT TGCAAGAGGG ATTGCGTGGC GTCTATCATT CTAGAGAGAT	60
AGACTTTGTA GAAAAAGTGG TTGTTTTAGA CAGCTGTCAA ATCCACCAAA AAGCGTTAAT	120
GCATTTGCAA GAAACTTTGA TGATAGAAGT GGATAGGCTT GATTTTTCTT TAGTGGAGCG	180
CTTGAACATT TTAGCGCGCA TGGAGA ATG AAA AGC ATG CGT TTT AGT TAC ATT	233
Met Lys Ser Met Arg Phe Ser Tyr Ile	
1 5	
GAG CCA AGA GCG AAA TAC CTT ATC AGC AAG CTT TCT AAA ATT TGG GTT	281
Glu Pro Arg Ala Lys Tyr Leu Ile Ser Lys Leu Ser Lys Ile Trp Val	
10 15 20 25	
TTT TAC ATT TTT TTA TCT TTT GTG GTA ATA GGG GGG TTA GTG TGG TTT	329
Phe Tyr Ile Phe Leu Ser Phe Val Val Ile Gly Gly Leu Val Trp Phe	
30 35 40	
ATG CAC AAC GCC ATT AAA AGC ACT CAA GAC AAC GCG TCC AGT TTG ACG	377
Met His Asn Ala Ile Lys Ser Thr Gln Asp Asn Ala Ser Ser Leu Thr	
45 50 55	

-199-

ATC CAA GAA AGG CTC TAC CGC CAT GAA ATC AGC CGC TTA CAG GTT AAG	425
Ile Gln Glu Arg Leu Tyr Arg His Glu Ile Ser Arg Leu Gln Val Lys	
60 65 70	
ACT GAT GAA ACC TTA AAA CTC ATT AAA GAA GCC AAA AAG CGT TTG AAT	473
Thr Asp Glu Thr Leu Lys Leu Ile Lys Glu Ala Lys Lys Arg Leu Asn	
75 80 85	
TAT AAC GAT GAT ATA CGA GAT GTT TTG CAA GGG CTT TTG AAT ATT GTG	521
Tyr Asn Asp Asp Ile Arg Asp Val Leu Gln Gly Leu Leu Asn Ile Val	
90 95 100 105	
CCG GAT TCC ATC ACT ATT AAT AGC ATT GAA ATA GAC CAG CAA AGC GTG	569
Pro Asp Ser Ile Thr Ile Asn Ser Ile Glu Ile Asp Gln Gln Ser Val	
110 115 120	
GTT GTT AGC GGT AAA ACC CCT TCT AAA GAA GCC TTT TAT TTT TTG TTT	617
Val Val Ser Gly Lys Thr Pro Ser Lys Glu Ala Phe Tyr Phe Leu Phe	
125 130 135	
CAA AAC AAA CTA AAC CCC ATG TTT GAT TAT TCT AGG GCG GAA TTT TTC	665
Gln Asn Lys Leu Asn Pro Met Phe Asp Tyr Ser Arg Ala Glu Phe Phe	
140 145 150	
CCC TTA AGC GAT GGG TGG TTT AAT TTT GTC TCC ACC AAC TTT TCT AAT	713
Pro Leu Ser Asp Gly Trp Phe Asn Phe Val Ser Thr Asn Phe Ser Asn	
155 160 165	
TCC TTA CTG ATA AAA AAT CCG GAG TCT ATT AAA TGAAGCCATT GCATTTTTTCA	766
Ser Leu Leu Ile Lys Asn Pro Glu Ser Ile Lys	
170 175 180	
CACCTGGACA GAGAGCAATC AGGCGATGTG GGGTTTATCA TTAAAAACCT CGTTTTTTTTA	826
GGGGTTTTTT CTTATTGGG TTGGTTGAAT ACCGAGTATT TTCTATGGCC TAGCATGCTG	886
GAATTAAAAA AAATCCTTTT AGAAGAAAAT CGTAAAAAAA	926

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Lys Ser Met Arg Phe Ser Tyr Ile Glu Pro Arg Ala Lys Tyr Leu	
1 5 10 15	
Ile Ser Lys Leu Ser Lys Ile Trp Val Phe Tyr Ile Phe Leu Ser Phe	
20 25 30	
Val Val Ile Gly Gly Leu Val Trp Phe Met His Asn Ala Ile Lys Ser	
35 40 45	
Thr Gln Asp Asn Ala Ser Ser Leu Thr Ile Gln Glu Arg Leu Tyr Arg	

-200-

50	55	60
His Glu Ile Ser Arg	Leu Gln Val Lys Thr Asp	Glu Thr Leu Lys Leu
65	70	75
Ile Lys Glu Ala Lys	Lys Arg Leu Asn Tyr Asn	Asp Asp Ile Arg Asp
	85	90
Val Leu Gln Gly Leu	Leu Asn Ile Val Pro	Asp Ser Ile Thr Ile Asn
	100	105
Ser Ile Glu Ile Asp	Gln Gln Ser Val Val	Val Ser Gly Lys Thr Pro
	115	120
Ser Lys Glu Ala Phe	Tyr Phe Leu Phe Gln	Asn Lys Leu Asn Pro Met
	130	135
Phe Asp Tyr Ser Arg	Ala Glu Phe Phe Pro	Leu Ser Asp Gly Trp Phe
145	150	155
Asn Phe Val Ser Thr	Asn Phe Ser Asn Ser	Leu Leu Ile Lys Asn Pro
	165	170
Glu Ser Ile Lys		175
	180	

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1440 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 151...1299
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

AGCACTTTTCG CTTTTCATTG TTTTGATGCG ACTTCTAGTT TCAGGCTTTT GCAAGTGTTA	60
AACGATGAGG TGAGCGATGC GTTTTAAATC ATACAAGATT TTAAAGAACA GCGCATCATT	120
CATAAAATCA TTCAAACCCA TTTCAAACGC ATG TGC GTG GTT TTG AGC GTG AAA	174
Met Cys Val Val Leu Ser Val Lys	
1 5	
AGA GAT GGT GAA AAA ACT TTA GAA AAT AAT GAA GAA AAT AAA GAT GAA	222
Arg Asp Gly Glu Lys Thr Leu Glu Asn Asn Glu Glu Asn Lys Asp Glu	
10 15 20	
AAG CTT ATT TTG ATT GAT GAA TTT GAA GTT TTA GCC AAT AAA TTC ATT	270
Lys Leu Ile Leu Ile Asp Glu Phe Glu Val Leu Ala Asn Lys Phe Ile	
25 30 35 40	
TCT CGT TTG CCC AAT ATC CCT AGC ACC CCT AGA GAG TTT GGG TTA GGC	318
Ser Arg Leu Pro Asn Ile Pro Ser Thr Pro Arg Glu Phe Gly Leu Gly	
45 50 55	
AAG GGC GAG ATC ATG GAG ATT GAT GTG CCT TTT GGG AGT ATT TTT GCT	366
Lys Gly Glu Ile Met Glu Ile Asp Val Pro Phe Gly Ser Ile Phe Ala	

-201-

60	65	70	
TAC AGA CAC ATT GGC TCT ATC AGA CAA AAA GAA TAC AGG ATT GTA GGG Tyr Arg His Ile Gly Ser Ile Arg Gln Lys Glu Tyr Arg Ile Val Gly 75 80 85			414
CTT TAT CGC AAC GAT GTT TTG TTG CTC TCC ACT AAA TCT TTA GTT ATC Leu Tyr Arg Asn Asp Val Leu Leu Leu Ser Thr Lys Ser Leu Val Ile 90 95 100			462
CAG CCG CGA GAC ATT CTC TTA GTG GCG GGT AAT CCG GAA ATT TTG AAT Gln Pro Arg Asp Ile Leu Leu Val Ala Gly Asn Pro Glu Ile Leu Asn 105 110 115 120			510
GCG GTG TAT CTT CAA GTC AAA AGC AAT GTG GGG CAG TTC CCA GCC CCC Ala Val Tyr Leu Gln Val Lys Ser Asn Val Gly Gln Phe Pro Ala Pro 125 130 135			558
TTT GGT AAG AGC ATT TAT TTA TAC ATT GAT ATG CGT TTG CAG AAC AGA Phe Gly Lys Ser Ile Tyr Leu Tyr Ile Asp Met Arg Leu Gln Asn Arg 140 145 150			606
AAA GCG ATG ATG CGC GAT GTG TAT CAA GCC TTG TTT TTG CAC AAA CAT Lys Ala Met Met Arg Asp Val Tyr Gln Ala Leu Phe Leu His Lys His 155 160 165			654
TTA AAG AGC TAC AAG CTC TAC ATT CAG GTT TTA CAC CCC ACT AGC CCT Leu Lys Ser Tyr Lys Leu Tyr Ile Gln Val Leu His Pro Thr Ser Pro 170 175 180			702
AAG TTT TAC CAT AAA TTT TTA GCG CTA GAA ACC GAA AGC ATT GAA GTG Lys Phe Tyr His Lys Phe Leu Ala Leu Glu Thr Glu Ser Ile Glu Val 185 190 195 200			750
AAT TTT GAT TTT TAC AGG AAA AGT TTT ATC CAA AAA CTC CAT GAA GAC Asn Phe Asp Phe Tyr Arg Lys Ser Phe Ile Gln Lys Leu His Glu Asp 205 210 215			798
CAC CAG AAA AAA ATG GGC CTA ATC GTG GTA GGC AGA GAG CTT TTT TTA His Gln Lys Lys Met Gly Leu Ile Val Val Gly Arg Glu Leu Phe Leu 220 225 230			846
TCT AAA AAA CAC CGA AAG GCC TTG TAT AAA ACA GCC ACC CCA GTT TAT Ser Lys Lys His Arg Lys Ala Leu Tyr Lys Thr Ala Thr Pro Val Tyr 235 240 245			894
AAA ACC AAC ACT TCT GGC TTG TCT AAA ACC TCT CAA AGC GTG GTG GTA Lys Thr Asn Thr Ser Gly Leu Ser Lys Thr Ser Gln Ser Val Val Val 250 255 260			942
TTG AAT GAA AGT TTG GAT ATT AAT GAG GAC ATG TCT TCA GTG ATT TTT Leu Asn Glu Ser Leu Asp Ile Asn Glu Asp Met Ser Ser Val Ile Phe 265 270 275 280			990
GAT GTG TCT ATG CAA ATG GAT TTG GGC TTG TTG CTC TAT GAT TTT GAC			1038

-202-

Asp Val Ser Met Gln Met Asp Leu Gly Leu Leu Leu Tyr Asp Phe Asp	
285 290 295	
CCT AAC AAG CGC TAT AAA AAC GAG ATT GTC AAT CAT TAT GAA AAT TTA	1086
Pro Asn Lys Arg Tyr Lys Asn Glu Ile Val Asn His Tyr Glu Asn Leu	
300 305 310	
GCC AAC GCG TTC AAC CGC AAG ATT GAG ATT TTC CAA ACC GAT ATT AGA	1134
Ala Asn Ala Phe Asn Arg Lys Ile Glu Ile Phe Gln Thr Asp Ile Arg	
315 320 325	
AAT CCT ATC ATG TAT CTC AAT TCT TTA AGA AAT CCC ATT TTG CAT TTC	1182
Asn Pro Ile Met Tyr Leu Asn Ser Leu Arg Asn Pro Ile Leu His Phe	
330 335 340	
ATG CCT TTT GAA GAG TGC ATC ACG CAC ACG CGC TTT TGG TGG TTT TTA	1230
Met Pro Phe Glu Glu Cys Ile Thr His Thr Arg Phe Trp Trp Phe Leu	
345 350 355 360	
TCC ACT AAA GTG GAA AAA TTA GCG TTT TTA AAC GAT GAT AAC CCT CAA	1278
Ser Thr Lys Val Glu Lys Leu Ala Phe Leu Asn Asp Asp Asn Pro Gln	
365 370 375	
ATT TTT ATC CCT GTA GCG GAG TGAAAGAATG CAAGAAATTT TAATCCCTTT AAAA	1333
Ile Phe Ile Pro Val Ala Glu	
380	
GAAAAAACT ATAAAGTGTT TTTGGGGGAA CTGCCTGAAA TAAAATTGAA AAAAAAGCC	1393
CTCATCATTA GCGATAGCAT CGTAGCCGGG TTGCATTTGC CCTATTT	1440

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 383 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Met Cys Val Val Leu Ser Val Lys Arg Asp Gly Glu Lys Thr Leu Glu	
1 5 10 15	
Asn Asn Glu Glu Asn Lys Asp Glu Lys Leu Ile Leu Ile Asp Glu Phe	
20 25 30	
Glu Val Leu Ala Asn Lys Phe Ile Ser Arg Leu Pro Asn Ile Pro Ser	
35 40 45	
Thr Pro Arg Glu Phe Gly Leu Gly Lys Gly Glu Ile Met Glu Ile Asp	
50 55 60	
Val Pro Phe Gly Ser Ile Phe Ala Tyr Arg His Ile Gly Ser Ile Arg	
65 70 75 80	
Gln Lys Glu Tyr Arg Ile Val Gly Leu Tyr Arg Asn Asp Val Leu Leu	
85 90 95	
Leu Ser Thr Lys Ser Leu Val Ile Gln Pro Arg Asp Ile Leu Leu Val	

- 203 -

				100				105					110		
Ala	Gly	Asn	Pro	Glu	Ile	Leu	Asn	Ala	Val	Tyr	Leu	Gln	Val	Lys	Ser
		115					120					125			
Asn	Val	Gly	Gln	Phe	Pro	Ala	Pro	Phe	Gly	Lys	Ser	Ile	Tyr	Leu	Tyr
		130				135					140				
Ile	Asp	Met	Arg	Leu	Gln	Asn	Arg	Lys	Ala	Met	Met	Arg	Asp	Val	Tyr
145					150					155					160
Gln	Ala	Leu	Phe	Leu	His	Lys	His	Leu	Lys	Ser	Tyr	Lys	Leu	Tyr	Ile
				165					170					175	
Gln	Val	Leu	His	Pro	Thr	Ser	Pro	Lys	Phe	Tyr	His	Lys	Phe	Leu	Ala
			180					185					190		
Leu	Glu	Thr	Glu	Ser	Ile	Glu	Val	Asn	Phe	Asp	Phe	Tyr	Arg	Lys	Ser
		195					200					205			
Phe	Ile	Gln	Lys	Leu	His	Glu	Asp	His	Gln	Lys	Lys	Met	Gly	Leu	Ile
	210					215					220				
Val	Val	Gly	Arg	Glu	Leu	Phe	Leu	Ser	Lys	Lys	His	Arg	Lys	Ala	Leu
225					230					235					240
Tyr	Lys	Thr	Ala	Thr	Pro	Val	Tyr	Lys	Thr	Asn	Thr	Ser	Gly	Leu	Ser
				245					250					255	
Lys	Thr	Ser	Gln	Ser	Val	Val	Val	Leu	Asn	Glu	Ser	Leu	Asp	Ile	Asn
			260					265					270		
Glu	Asp	Met	Ser	Ser	Val	Ile	Phe	Asp	Val	Ser	Met	Gln	Met	Asp	Leu
		275					280					285			
Gly	Leu	Leu	Leu	Tyr	Asp	Phe	Asp	Pro	Asn	Lys	Arg	Tyr	Lys	Asn	Glu
	290					295					300				
Ile	Val	Asn	His	Tyr	Glu	Asn	Leu	Ala	Asn	Ala	Phe	Asn	Arg	Lys	Ile
305					310					315					320
Glu	Ile	Phe	Gln	Thr	Asp	Ile	Arg	Asn	Pro	Ile	Met	Tyr	Leu	Asn	Ser
				325					330					335	
Leu	Arg	Asn	Pro	Ile	Leu	His	Phe	Met	Pro	Phe	Glu	Glu	Cys	Ile	Thr
			340					345					350		
His	Thr	Arg	Phe	Trp	Trp	Phe	Leu	Ser	Thr	Lys	Val	Glu	Lys	Leu	Ala
		355					360					365			
Phe	Leu	Asn	Asp	Asp	Asn	Pro	Gln	Ile	Phe	Ile	Pro	Val	Ala	Glu	
	370					375					380				

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 517 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 51...464
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

AGATTTTCATT CGAGGTAGAA AATACATTGA AAAAGCGTGT GAATTAAACG ATG GTA

-205-

```

      1             5             10             15
Tyr Tyr Val Lys Ala Cys Glu Leu Asn Glu Met Phe Gly Cys Leu Ser
      20             25             30
Leu Val Ser Asn Ser Gln Ile Asn Lys Gln Lys Leu Phe Gln Tyr Leu
      35             40             45
Ser Lys Ala Cys Glu Leu Asn Ser Gly Asn Gly Cys Arg Phe Leu Gly
      50             55             60
Asp Phe Tyr Glu Asn Gly Lys Tyr Val Lys Lys Asp Leu Arg Lys Ala
      65             70             75             80
Ala Gln Tyr Tyr Ser Lys Ala Cys Gly Leu Asn Asp Gln Asp Gly Cys
      85             90             95
Leu Ile Leu Gly Tyr Lys Gln Tyr Ala Gly Lys Gly Val Val Lys Asn
      100            105            110
Glu Lys Gln Ala Val Lys Thr Phe Glu Lys Ala Cys Arg Leu Gly Ser
      115            120            125
Glu Asp Ala Cys Gly Ile Leu Asn Asn Tyr
      130            135

```

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1663 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 68...1600
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```

AAATGTTAGA AACCTTACA AAACAAGCTA ATATATTCTA TTCAATTGTC CTCAAGGACA      60
AACAAAC  ATG AAA AAA CTT CTT TAT ACC ATA CTC GCG CTT CTT TTA ATC      109
      Met Lys Lys Leu Leu Tyr Thr Ile Leu Ala Leu Leu Ile
      1             5             10

GGC CTT TTA ACA ATC TAT CTC ATC CTT TTT ACA GAA TGG GGG AAT AAG      157
Gly Leu Leu Thr Ile Tyr Leu Ile Leu Phe Thr Glu Trp Gly Asn Lys
      15             20             25             30

ATC ATC GCT TCG TAT ATA GAG AAA AAA ATC AAC CCG AAC GAG CAC TAC      205
Ile Ile Ala Ser Tyr Ile Glu Lys Lys Ile Asn Pro Asn Glu His Tyr
      35             40             45

TTG AGC GTT AAA ACC TTT AAA TTG AGA TTC AAC TCT TTG GAT TTT AAA      253
Leu Ser Val Lys Thr Phe Lys Leu Arg Phe Asn Ser Leu Asp Phe Lys
      50             55             60

GCT CAA GCC AAC GAT GAT TCC ACG CTC ATT CTT AAG GGG GAT TTT TCA      301
Ala Gln Ala Asn Asp Asp Ser Thr Leu Ile Leu Lys Gly Asp Phe Ser

```

-206-

65	70	75	
CTT TTA AAG CAA AGC GTA AAT TTG AAT TAC CAT ATA GAT ATT AAA GAT			349
Leu Leu Lys Gln Ser Val Asn Leu Asn Tyr His Ile Asp Ile Lys Asp			
80	85	90	
TTA CGC TCT TTC AAA GAA TGG ATA CCC TAC CCT TTA AGG GGG GCT GTT			397
Leu Arg Ser Phe Lys Glu Trp Ile Pro Tyr Pro Leu Arg Gly Ala Val			
95	100	105	110
ATC ACT TCT GGG AAT ATT AAA GGG CAT AGA AAA GCC CTT ATG ATT CAA			445
Ile Thr Ser Gly Asn Ile Lys Gly His Arg Lys Ala Leu Met Ile Gln			
115	120	125	
GGC GTC TCT AAT GTG GCT CAA TCC CAC ACT GCC TAC AAT GCC CTT TTA			493
Gly Val Ser Asn Val Ala Gln Ser His Thr Ala Tyr Asn Ala Leu Leu			
130	135	140	
GAT GAT TTC AAG CTT TCT CGC TTA AAT TTG AAC GCA CAA GAC GCC AAT			541
Asp Asp Phe Lys Leu Ser Arg Leu Asn Leu Asn Ala Gln Asp Ala Asn			
145	150	155	
TTA GAA GAT TTG CTT TAT TTA ATC AAT CGC CCC GCT TAT GCG AAC GCA			589
Leu Glu Asp Leu Leu Tyr Leu Ile Asn Arg Pro Ala Tyr Ala Asn Ala			
160	165	170	
AAA GTG TCC TTA CAG GCG GAT TTT AAC TCT CTA AAG CCT TTA GAG GGG			637
Lys Val Ser Leu Gln Ala Asp Phe Asn Ser Leu Lys Pro Leu Glu Gly			
175	180	185	190
CAT TTG ATC CTA ACA GCT AAT AAC GCT TTA ATC AAT AAC GCC CTA ATC			685
His Leu Ile Leu Thr Ala Asn Asn Ala Leu Ile Asn Asn Ala Leu Ile			
195	200	205	
AAT CAA ATT TTT CAT TTA AAC CTT AAA GAC ACG CTT GTT TTC AGC CTC			733
Asn Gln Ile Phe His Leu Asn Leu Lys Asp Thr Leu Val Phe Ser Leu			
210	215	220	
TCG CAT TCA AGC GAC TTT AAA GGA AAC AAA GCC ATC AGC GAT ACC ACC			781
Ser His Ser Ser Asp Phe Lys Gly Asn Lys Ala Ile Ser Asp Thr Thr			
225	230	235	
CTG ACT AGC CCT TTA GCC AAT TTC AAA GCC CTA AAA AGC GAA TAC CTT			829
Leu Thr Ser Pro Leu Ala Asn Phe Lys Ala Leu Lys Ser Glu Tyr Leu			
240	245	250	
TTC TCT ATT TTA AAA CTC AAC GCC CCC TAC ACT TTA GAA ATC CCC AAT			877
Phe Ser Ile Leu Lys Leu Asn Ala Pro Tyr Thr Leu Glu Ile Pro Asn			
255	260	265	270
CTA GCC AAA CTC TAT AAC ATT ACC AAC CAC CCC TTA AAA GGG AGC TTG			925
Leu Ala Lys Leu Tyr Asn Ile Thr Asn His Pro Leu Lys Gly Ser Leu			
275	280	285	
ACT TTA AAA GGC GCT ATA GAA CAA AGC CCC AAA CTT TTA AAA GTC AGC			973

-207-

Thr	Leu	Lys	Gly	Ala	Ile	Glu	Gln	Ser	Pro	Lys	Leu	Leu	Lys	Val	Ser		
			290					295					300				
GGC	CAT	TCA	AAT	TTA	CTA	GAC	GGC	GCG	CTG	GAT	TTC	ACG	CTT	TTA	AAT	1021	
Gly	His	Ser	Asn	Leu	Leu	Asp	Gly	Ala	Leu	Asp	Phe	Thr	Leu	Leu	Asn		
		305					310				315						
AAA	GAT	TTG	AAA	GGG	CGT	TTT	TCC	AAT	ATT	TCC	ACT	TTA	AAA	GCT	TTA	1069	
Lys	Asp	Leu	Lys	Gly	Arg	Phe	Ser	Asn	Ile	Ser	Thr	Leu	Lys	Ala	Leu		
	320					325					330						
GAT	TTA	TTC	CAT	TAC	CCT	AAG	TTT	TTC	CAA	TCC	GTT	GCA	GAC	GCT	AAT	1117	
Asp	Leu	Phe	His	Tyr	Pro	Lys	Phe	Phe	Gln	Ser	Val	Ala	Asp	Ala	Asn		
	335				340					345					350		
TTG	GAT	TAT	GAT	CTT	ATC	GCT	AAG	CAA	GGC	GTA	TTG	AAA	GCC	CGC	CTA	1165	
Leu	Asp	Tyr	Asp	Leu	Ile	Ala	Lys	Gln	Gly	Val	Leu	Lys	Ala	Arg	Leu		
				355				360						365			
AAA	AAC	GCA	AGA	TTC	CTC	AAA	AAT	GCA	TTC	AGC	GAT	TTT	CTC	TAC	TCC	1213	
Lys	Asn	Ala	Arg	Phe	Leu	Lys	Asn	Ala	Phe	Ser	Asp	Phe	Leu	Tyr	Ser		
		370						375					380				
ATT	TCT	AAA	TTT	GAT	ATT	ACA	AAA	GAA	ATT	TAT	AAC	GAT	GCC	AAT	CTG	1261	
Ile	Ser	Lys	Phe	Asp	Ile	Thr	Lys	Glu	Ile	Tyr	Asn	Asp	Ala	Asn	Leu		
		385					390					395					
GTA	AGC	CAA	ATC	AAC	CAG	CAA	CGC	CTG	CTC	TCT	GAT	CTG	AGT	TTA	AAA	1309	
Val	Ser	Gln	Ile	Asn	Gln	Gln	Arg	Leu	Leu	Ser	Asp	Leu	Ser	Leu	Lys		
	400					405					410						
AGC	CCC	AAA	ACC	CAA	TTG	AAA	ATC	CAT	AAC	GGT	TTG	TTG	GAT	TTA	AAC	1357	
Ser	Pro	Lys	Thr	Gln	Leu	Lys	Ile	His	Asn	Gly	Leu	Leu	Asp	Leu	Asn		
	415				420					425					430		
ACC	AAA	CAA	ATG	AAC	ATG	CTC	ATG	GAT	GCG	GAA	ATT	TTA	AAA	TTC	ATT	1405	
Thr	Lys	Gln	Met	Asn	Met	Leu	Met	Asp	Ala	Glu	Ile	Leu	Lys	Phe	Ile		
				435				440						445			
TTT	AAA	ATG	AAA	CTT	CAA	GGC	AAC	ATG	CAC	CAG	CCA	AAA	TTT	TCT	CTC	1453	
Phe	Lys	Met	Lys	Leu	Gln	Gly	Asn	Met	His	Gln	Pro	Lys	Phe	Ser	Leu		
			450				455						460				
ATT	TTA	AAC	GAA	AAA	GCC	ATT	CAG	CAA	AAC	TTG	CAA	CAA	GGC	TTG	AAA	1501	
Ile	Leu	Asn	Glu	Lys	Ala	Ile	Gln	Gln	Asn	Leu	Gln	Gln	Gly	Leu	Lys		
		465				470					475						
GAA	ATC	TTA	AAA	AAC	GAC	ACC	CTT	AAA	AAA	GGT	TTA	GAT	CAT	TTG	CTT	1549	
Glu	Ile	Leu	Lys	Asn	Asp	Thr	Leu	Lys	Lys	Gly	Leu	Asp	His	Leu	Leu		
	480					485				490							
AAA	GAT	GAT	AAG	CTC	AAA	GAA	AAG	CTT	GAA	AAA	GGG	CTT	AAG	GGG	CTT	1597	
Lys	Asp	Asp	Lys	Leu	Lys	Glu	Lys	Leu	Glu	Lys	Gly	Leu	Lys	Gly	Leu		
	495				500					505					510		
TTT	TAAAAATTTT	AAAGGATAGA	AATGGCGCAC	ATTTTAGTTA	GCGGGGCGAC	TTCAGG	1656										

Phe

GTTTGGA

1663

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 511 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```

Met Lys Lys Leu Leu Tyr Thr Ile Leu Ala Leu Leu Leu Ile Gly Leu
 1           5           10           15
Leu Thr Ile Tyr Leu Ile Leu Phe Thr Glu Trp Gly Asn Lys Ile Ile
          20           25           30
Ala Ser Tyr Ile Glu Lys Lys Ile Asn Pro Asn Glu His Tyr Leu Ser
          35           40           45
Val Lys Thr Phe Lys Leu Arg Phe Asn Ser Leu Asp Phe Lys Ala Gln
          50           55           60
Ala Asn Asp Asp Ser Thr Leu Ile Leu Lys Gly Asp Phe Ser Leu Leu
          65           70           75           80
Lys Gln Ser Val Asn Leu Asn Tyr His Ile Asp Ile Lys Asp Leu Arg
          85           90           95
Ser Phe Lys Glu Trp Ile Pro Tyr Pro Leu Arg Gly Ala Val Ile Thr
          100          105          110
Ser Gly Asn Ile Lys Gly His Arg Lys Ala Leu Met Ile Gln Gly Val
          115          120          125
Ser Asn Val Ala Gln Ser His Thr Ala Tyr Asn Ala Leu Leu Asp Asp
          130          135          140
Phe Lys Leu Ser Arg Leu Asn Leu Asn Ala Gln Asp Ala Asn Leu Glu
          145          150          155          160
Asp Leu Leu Tyr Leu Ile Asn Arg Pro Ala Tyr Ala Asn Ala Lys Val
          165          170          175
Ser Leu Gln Ala Asp Phe Asn Ser Leu Lys Pro Leu Glu Gly His Leu
          180          185          190
Ile Leu Thr Ala Asn Asn Ala Leu Ile Asn Asn Ala Leu Ile Asn Gln
          195          200          205
Ile Phe His Leu Asn Leu Lys Asp Thr Leu Val Phe Ser Leu Ser His
          210          215          220
Ser Ser Asp Phe Lys Gly Asn Lys Ala Ile Ser Asp Thr Thr Leu Thr
          225          230          235          240
Ser Pro Leu Ala Asn Phe Lys Ala Leu Lys Ser Glu Tyr Leu Phe Ser
          245          250          255
Ile Leu Lys Leu Asn Ala Pro Tyr Thr Leu Glu Ile Pro Asn Leu Ala
          260          265          270
Lys Leu Tyr Asn Ile Thr Asn His Pro Leu Lys Gly Ser Leu Thr Leu
          275          280          285
Lys Gly Ala Ile Glu Gln Ser Pro Lys Leu Leu Lys Val Ser Gly His
          290          295          300

```

-209-

```

Ser Asn Leu Leu Asp Gly -Ala Leu Asp Phe Thr Leu Leu Asn Lys Asp
305                      310                      315                      320
Leu Lys Gly Arg Phe Ser Asn Ile Ser Thr Leu Lys Ala Leu Asp Leu
                      325                      330                      335
Phe His Tyr Pro Lys Phe Phe Gln Ser Val Ala Asp Ala Asn Leu Asp
                      340                      345                      350
Tyr Asp Leu Ile Ala Lys Gln Gly Val Leu Lys Ala Arg Leu Lys Asn
                      355                      360                      365
Ala Arg Phe Leu Lys Asn Ala Phe Ser Asp Phe Leu Tyr Ser Ile Ser
                      370                      375                      380
Lys Phe Asp Ile Thr Lys Glu Ile Tyr Asn Asp Ala Asn Leu Val Ser
385                      390                      395                      400
Gln Ile Asn Gln Gln Arg Leu Leu Ser Asp Leu Ser Leu Lys Ser Pro
                      405                      410                      415
Lys Thr Gln Leu Lys Ile His Asn Gly Leu Leu Asp Leu Asn Thr Lys
                      420                      425                      430
Gln Met Asn Met Leu Met Asp Ala Glu Ile Leu Lys Phe Ile Phe Lys
                      435                      440                      445
Met Lys Leu Gln Gly Asn Met His Gln Pro Lys Phe Ser Leu Ile Leu
                      450                      455                      460
Asn Glu Lys Ala Ile Gln Gln Asn Leu Gln Gln Gly Leu Lys Glu Ile
465                      470                      475                      480
Leu Lys Asn Asp Thr Leu Lys Lys Gly Leu Asp His Leu Leu Lys Asp
                      485                      490                      495
Asp Lys Leu Lys Glu Lys Leu Glu Lys Gly Leu Lys Gly Leu Phe
                      500                      505                      510

```

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 947 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 292...645
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```

AGTGCATAAA CGCACAGACC CCAAAAAATGA AAGCTATTTT TGGCTAGGGC TACACCCTTT      60
AGAATGGCAA AAGCGCGAAA ATGAAGACAG ACTCTCTGAT TTTGACGCTA TTGCTTCAAA      120
CCATGCCTCT ATCACGCTT TAAATTTAGA CTTAACCAGT TATGATGATT TGAAAAGTTT      180
GGAATCTTGG CATGAGGGAA TGTTAAAGTG AGTAAAAAGC ACCGCTTGGC TTTTITAGGG      240
CTAATTGTTG GGGTTCATT CTTCTTTAGT GCGTGTGAGC ACCGCCTGCA C ATG GGG      297
                                         Met Gly
                                         1

TAT TAT TCA GAA GTT ACA GGG GAT TAT TTG TTC AAT TAT AAT TCC ACT      345
Tyr Tyr Ser Glu Val Thr Gly Asp Tyr Leu Phe Asn Tyr Asn Ser Thr

```

-210-

5	10	15	
ATC GTG GTG GCT TAT GAC AGA AGC GAT GCG ATG ACT TCT TAT TAT ATC	393		
Ile Val Val Ala Tyr Asp Arg Ser Asp Ala Met Thr Ser Tyr Tyr Ile			
20 25 30			
AAT GTG ATT GTT TAT GAA TTG CAA AAA TTA GGC TTT TAC AAT GTC TTC	441		
Asn Val Ile Val Tyr Glu Leu Gln Lys Leu Gly Phe Tyr Asn Val Phe			
35 40 45 50			
ACG CAA GCG GAA TTC CCA CTA GAT AAA GCC AAA AAT GTG ATC TAT GCG	489		
Thr Gln Ala Glu Phe Pro Leu Asp Lys Ala Lys Asn Val Ile Tyr Ala			
55 60 65			
CGC ATT GTC CGT AAC ATC TCA GCT GTG CCG TTC TAC CAA TAC AAT TAC	537		
Arg Ile Val Arg Asn Ile Ser Ala Val Pro Phe Tyr Gln Tyr Asn Tyr			
70 75 80			
CAA CTG ATT GAT CAA GTC AAT AAG CCT TGT TAT TTT CTT GGG GGG CAG	585		
Gln Leu Ile Asp Gln Val Asn Lys Pro Cys Tyr Phe Leu Gly Gly Gln			
85 90 95			
TTT TAT TGC TCT CAA ACC CTA CGG ATT ATT ACG CTA TCA ATG GCT TTA	633		
Phe Tyr Cys Ser Gln Thr Leu Arg Ile Ile Thr Leu Ser Met Ala Leu			
100 105 110			
GCG AGC AAA TTT TAATGAGTGC TAATTCGCAT TTTATTTTAG ATTGGTATGA TGTGG	690		
Ala Ser Lys Phe			
115			
TGTTGCAAAA ACGGGTTTTA TATGTGGATG GGAGCGTGAG CGGGAGGACT TGCGGCTATC	750		
AGATGCTGTA TAGGGATTG ATTAAGAGCA CGATCAAACG CATTGATTTT AACCGCCCTG	810		
AACGCTACTA CTACAATTTA AGACTGCCCC TTTATCAGCC ATGTTATAGG CAATGAAATG	870		
GTTATCAGGC GATTGTATCA ATTTTGCGCT AGCCATGTGG TGCGCAATTG CTCTTCTTTA	930		
AAATGCGCTC AAAATAT	947		

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Met Gly Tyr Tyr Ser Glu Val Thr Gly Asp Tyr Leu Phe Asn Tyr Asn	
1 5 10 15	
Ser Thr Ile Val Val Ala Tyr Asp Arg Ser Asp Ala Met Thr Ser Tyr	
20 25 30	
Tyr Ile Asn Val Ile Val Tyr Glu Leu Gln Lys Leu Gly Phe Tyr Asn	
35 40 45	
Val Phe Thr Gln Ala Glu Phe Pro Leu Asp Lys Ala Lys Asn Val Ile	

-211-

50		- 55		60											
Tyr	Ala	Arg	Ile	Val	Arg	Asn	Ile	Ser	Ala	Val	Pro	Phe	Tyr	Gln	Tyr
65					70				75					80	
Asn	Tyr	Gln	Leu	Ile	Asp	Gln	Val	Asn	Lys	Pro	Cys	Tyr	Phe	Leu	Gly
			85					90						95	
Gly	Gln	Phe	Tyr	Cys	Ser	Gln	Thr	Leu	Arg	Ile	Ile	Thr	Leu	Ser	Met
			100					105						110	
Ala	Leu	Ala	Ser	Lys	Phe										
			115												

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 875 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 348...716
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

TGCGGAGGGA	ATGTCTATGA	TAAAATCTCA	GAAAAATTTG	TAGAAAAAGT	GGATAACGGG	60
TTTTGAAAAT	TTTAATCCTT	TTTTTTATCT	GTTTAAACGC	ATTGTTTCGCC	CTAGATTCAA	120
ACGCACTTAA	AGCAGAGATT	AAAGAAGTTT	ACCTTAAAGA	ATACAAAGAC	TTAAAATTAG	180
AAATTGAAAC	CATTAACCTA	GAAATCCCAG	AGCGCTTTTC	TAACGCTTCC	ATTTTAAGCT	240
ATGAATTAAA	CGCTTCCAAT	AAGCTTAAAA	AAGATGGGGT	CGTGTTTTTA	AGGTTGGAAA	300
ATGATCCTAA	TTTACGCCTA	CCGGTGC GTT	ATAGCGTGAT	AGGCAGC	ATG CAG GCT	356
				Met	Gln Ala	
				1		
TTT AAA AGC	GTT AGC GCG	ATT AAA AAA	GAT GAA AAC	ATC ACC GCT	AAT	404
Phe Lys Ser	Val Ser Ala	Ile Lys Lys	Asp Glu Asn	Ile Thr Ala	Asn	
5		10		15		
AAC ACT CAA	AAA GAG CGC	ATT TTG TTT	GGT GCG CTT	TCT AAC CCC	TTA	452
Asn Thr Gln	Lys Glu Arg	Ile Leu Phe	Gly Ala Leu	Ser Asn Pro	Leu	
20		25		30	35	
TTA GAG GGC	GCG ATT GAT	AAA GTG AGC	GCG AAA AAT	TTT ATC CCC	CCT	500
Leu Glu Gly	Ala Ile Asp	Lys Val Ser	Ala Lys Asn	Phe Ile Pro	Pro	
	40		45		50	
AAC ACG CTT	TTA AGC ACG	GAT AAA ACC	CAA GCT TTA	ATT ATC GTG	CGT	548
Asn Thr Leu	Leu Ser Thr	Asp Lys Thr	Gln Ala Leu	Ile Ile Val	Arg	
	55		60		65	
AAA AAT GAC	ATT ATC ACC	GGG GTG TAT	GAA GAG GGG	CAA ATC AGC	ATA	596
Lys Asn Asp	Ile Ile Thr	Gly Val Tyr	Glu Glu Gly	Gln Ile Ser	Ile	

-212-

70	75	80	
GAA ATA AGC CTA AAA GCC CTA GAA AAT GGC GCG CTT AAT CAA ATC ATT			644
Glu Ile Ser Leu Lys Ala Leu Glu Asn Gly Ala Leu Asn Gln Ile Ile			
85	90	95	
CAA GCG AAA AAT TTA GAA AGC AAT AAA ATA CTC AAA GCA AAA GTG TTG			692
Gln Ala Lys Asn Leu Glu Ser Asn Lys Ile Leu Lys Ala Lys Val Leu			
100	105	110	115
AGC AGC TCT AAA GCG CAA ATC TTA TAAAGGACAT TCATGAAATT GGTTTTAGGC			746
Ser Ser Ser Lys Ala Gln Ile Leu			
120			
ATCAGTGGAG CGAGCGGGAT ACCCCTAGCC TTGCGGTTTT TAGAAAAATT ACCCAAAGAA			806
ATTGAAGTTT TTGTCGTGGC GTCTAAAAAC GCGCATGTCG TGGCGTTAGA AGAATCTAAT			866
ATTAACCTT			875

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Met	Gln	Ala	Phe	Lys	Ser	Val	Ser	Ala	Ile	Lys	Lys	Asp	Glu	Asn	Ile
1				5					10					15	
Thr	Ala	Asn	Asn	Thr	Gln	Lys	Glu	Arg	Ile	Leu	Phe	Gly	Ala	Leu	Ser
		20						25					30		
Asn	Pro	Leu	Glu	Gly	Ala	Ile	Asp	Lys	Val	Ser	Ala	Lys	Asn	Phe	
		35				40				45					
Ile	Pro	Pro	Asn	Thr	Leu	Leu	Ser	Thr	Asp	Lys	Thr	Gln	Ala	Leu	Ile
		50				55				60					
Ile	Val	Arg	Lys	Asn	Asp	Ile	Ile	Thr	Gly	Val	Tyr	Glu	Glu	Gly	Gln
65				70					75					80	
Ile	Ser	Ile	Glu	Ile	Ser	Leu	Lys	Ala	Leu	Glu	Asn	Gly	Ala	Leu	Asn
		85						90					95		
Gln	Ile	Ile	Gln	Ala	Lys	Asn	Leu	Glu	Ser	Asn	Lys	Ile	Leu	Lys	Ala
		100					105						110		
Lys	Val	Leu	Ser	Ser	Ser	Lys	Ala	Gln	Ile	Leu					
		115					120								

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-213-

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 160...345
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```

GGCATCACTT TTAACATGAC CCCTTCTCCA GGCGCGACGA GTTGTTTGCA AAACGCCCTT      60
GTGGATTCCC AAGAAATCGC TCGTATTTG GGCGAGAGCT TTGAATTAGA ACGCTTTTAT      120
AAAGATTTAT CCCCAGAAGA ATTGGAAAAT TAAAAACGC ATG CAA AAA GAA CAA      174
                               Met Gln Lys Glu Gln
                               1           5

GAA GCC CAA GAA ATC GCT AAA AAA GCC GTT AAA ATC GTG TTT TTT TTA      222
Glu Ala Gln Glu Ile Ala Lys Lys Ala Val Lys Ile Val Phe Phe Leu
                               10           15           20

GGG CTT GTG GTG GTG CTT TTG ATG ATG ATA AAC CTT TAC ATG CTC ATC      270
Gly Leu Val Val Val Leu Leu Met Met Ile Asn Leu Tyr Met Leu Ile
                               25           30           35

AAT CAA ATC AAC GCG AGC GCT CAA ATG AGC CAC CAA ATC AAA AAG ATA      318
Asn Gln Ile Asn Ala Ser Ala Gln Met Ser His Gln Ile Lys Lys Ile
                               40           45           50

GAA GAA AGG CTT AAT CAG GAG CAA AAA TAAAAAAGGC TTTTGTGTAT TTTTACG      372
Glu Glu Arg Leu Asn Gln Glu Gln Lys
                               55           60

ATCAAATAGT AAAGAGCTTA TC      394

```

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

```

Met Gln Lys Glu Gln Glu Ala Gln Glu Ile Ala Lys Lys Ala Val Lys
1           5           10           15
Ile Val Phe Phe Leu Gly Leu Val Val Leu Leu Met Met Ile Asn
20           25           30
Leu Tyr Met Leu Ile Asn Gln Ile Asn Ala Ser Ala Gln Met Ser His
35           40           45
Gln Ile Lys Lys Ile Glu Glu Arg Leu Asn Gln Glu Gln Lys
50           55           60

```

-214-

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 982 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 320...880
- (D) OTHER INFORMATION:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 320...400
- (D) OTHER INFORMATION:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 401...880
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

```

AGATTAGCAG CAGCAGGGAT TTTTAAATTC CTGGCCAACA GGGGCGGTTG GAAAAAATA      60
CGATTAAAAA GGCAAACGCT TTGAAAGTAT TTTTTCATAG AAATTCCTT TTGTTAAATG      120
ATTGAAGTTG GTGATTATAC CTATTTGTAT CTTAAAAATT TGATTTTAAA AGTTTGAGAT      180
GGTTTTGTAG GTGTATCCCA CTTATCCAAT TTATATCAAT ATTTTCACTC TAAAACCTC      240
ATCCTTGATA AAAAATTAAA CCTTTTAGAA AAATAACCGA TTTTAGGGTG TAACTTTAAT      300
TCAACAAGAA GGATTTATT ATG ATT AAA AGA ATT GCT TGT ATT TTA AGC TTG      352
                Met Ile Lys Arg Ile Ala Cys Ile Leu Ser Leu
                -27   -25                               -20

AGT GCG AGT TTA GCG CTG GCT GGC GAA GTG AAT GGG TTT TTC ATG GGT      400
Ser Ala Ser Leu Ala Leu Ala Gly Glu Val Asn Gly Phe Phe Met Gly
    -15                -10                -5

GCG GGT TAT CAG CAA GGT CGT TAT GGT CCT TAT AAC AGC AAT TAC TCT      448
Ala Gly Tyr Gln Gln Gly Arg Tyr Gly Pro Tyr Asn Ser Asn Tyr Ser
    1                5                10                15

GAT TGG CGC CAT GGC AAT GAT CTT TAT GGT TTG AAT TTC AAA TTA GGT      496
Asp Trp Arg His Gly Asn Asp Leu Tyr Gly Leu Asn Phe Lys Leu Gly
                20                25                30

TTT GTA GGC TTT GCC AAT AAA TGG TTT GGG GCT AGG GTG TAT GGC TTT      544
Phe Val Gly Phe Ala Asn Lys Trp Phe Gly Ala Arg Val Tyr Gly Phe
    35                40                45

TTA GAT TGG TTT AAC ACT TCA GGG ACA GAA CAC ACC AAA ACC AAT TTG      592

```

-215-

Leu	Asp	Trp	Phe	Asn	Trp	Ser	Gly	Thr	Glu	His	Thr	Lys	Thr	Asn	Leu	
50						55					60					
CTC	ACC	TAT	GGT	GGC	GGT	GGC	GAT	TTG	ATT	GTC	AAT	CTC	ATT	CCT	TTG	640
Leu	Thr	Tyr	Gly	Gly	Gly	Gly	Asp	Leu	Ile	Val	Asn	Leu	Ile	Pro	Leu	
65					70					75					80	
GAT	AAA	TTC	GCT	CTA	GGT	CTC	ATC	GGT	GGC	GTT	CAA	TTA	GCC	GGA	AAC	688
Asp	Lys	Phe	Ala	Leu	Gly	Leu	Ile	Gly	Gly	Val	Gln	Leu	Ala	Gly	Asn	
				85					90					95		
ACT	TGG	ATG	TTC	CCT	TAT	GAT	GTC	AAT	CAA	ACG	AGA	TTC	CAG	TTC	TTA	736
Thr	Trp	Met	Phe	Pro	Tyr	Asp	Val	Asn	Gln	Thr	Arg	Phe	Gln	Phe	Leu	
			100					105					110			
TGG	AAT	TTA	GGC	GGA	AGA	ATG	CGT	GTT	GGG	GAT	CGC	AGT	GCG	TTT	GAA	784
Trp	Asn	Leu	Gly	Gly	Arg	Met	Arg	Val	Gly	Asp	Arg	Ser	Ala	Phe	Glu	
		115					120					125				
GCA	GGC	GTG	AAA	TTC	CCT	ATG	GTT	AAT	CAA	GGC	AAC	AAA	GAT	GTT	AGG	832
Ala	Gly	Val	Lys	Phe	Pro	Met	Val	Asn	Gln	Gly	Asn	Lys	Asp	Val	Arg	
	130						135				140					
GCT	TAT	CCG	CTA	CTA	TTC	TTG	GGT	ATG	TGG	ATT	ATG	TTC	TTC	ACT	TTC	881
Ala	Tyr	Pro	Leu	Leu	Phe	Leu	Gly	Met	Trp	Ile	Met	Phe	Phe	Thr	Phe	
145					150					155					160	
AATTTATTCC	TTTCATTTCG	TCTTCTTCAT	CAAATCAACC	CTAACCCACT	CTTAAAAGGT											941
TGGGGTTCAA	AAATCTTTTT	CATAAATAAA	ATTGCCTTA	A												982

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Met	Ile	Lys	Arg	Ile	Ala	Cys	Ile	Leu	Ser	Leu	Ser	Ala	Ser	Leu	Ala	
-27		-25					-20					-15				
Leu	Ala	Gly	Glu	Val	Asn	Gly	Phe	Phe	Met	Gly	Ala	Gly	Tyr	Gln	Gln	
-10					-5					1				5		
Gly	Arg	Tyr	Gly	Pro	Tyr	Asn	Ser	Asn	Tyr	Ser	Asp	Trp	Arg	His	Gly	
				10					15					20		
Asn	Asp	Leu	Tyr	Gly	Leu	Asn	Phe	Lys	Leu	Gly	Phe	Val	Gly	Phe	Ala	
		25						30				35				
Asn	Lys	Trp	Phe	Gly	Ala	Arg	Val	Tyr	Gly	Phe	Leu	Asp	Trp	Phe	Asn	
	40					45					50					
Thr	Ser	Gly	Thr	Glu	His	Thr	Lys	Thr	Asn	Leu	Leu	Thr	Tyr	Gly	Gly	
55						60					65					

-216-

Gly	Gly	Asp	Leu	Ile	Val	Asn	Leu	Ile	Pro	Leu	Asp	Lys	Phe	Ala	Leu
70					75				80					85	
Gly	Leu	Ile	Gly	Gly	Val	Gln	Leu	Ala	Gly	Asn	Thr	Trp	Met	Phe	Pro
			90						95					100	
Tyr	Asp	Val	Asn	Gln	Thr	Arg	Phe	Gln	Phe	Leu	Trp	Asn	Leu	Gly	Gly
			105					110					115		
Arg	Met	Arg	Val	Gly	Asp	Arg	Ser	Ala	Phe	Glu	Ala	Gly	Val	Lys	Phe
		120					125					130			
Pro	Met	Val	Asn	Gln	Gly	Asn	Lys	Asp	Val	Arg	Ala	Tyr	Pro	Leu	Leu
	135					140					145				
Phe	Leu	Gly	Met	Trp	Ile	Met	Phe	Phe	Thr	Phe					
150					155					160					

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 262...777
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CCAATGGAGG	CGTTTCCAAA	AACCCAAACG	GGCGCTTTTT	AAAGAAAAAT	CTCAAAAAAT	60
TCAGGGAGCA	AGCGGTAAAA	ATCGTAGAAA	AACGCTTGAT	AAAAGAGAAT	ATGCAACTGA	120
GCGATTTTAA	TGAAGAAGAA	TTAAAAATCA	TGTTTGAAGC	TGAAGAAAAA	AGGTTGTTAG	180
AGCAAATCCA	CGCTAAAGAA	TTGAAAGAAA	AGCAAGAAAA	AACCACCAAG	CATTTTAAAG	240
AAGTTTGGGA AAAGGGCGAA A ATG AGC AAG AAA AAT AGC GTA ATT TCT GGT						291
		Met	Ser	Lys	Lys	Asn
		1			5	10
TTA ATG AAT TTT TTT AGC GAA AAG AAT GAA CGC TGG CTC TTA GCC CAC						339
Leu	Met	Asn	Phe	Phe	Ser	Glu
		15			20	25
AGG CAC ACG AGA GGG TTT GTG ATA GTG GCG TGG CTT TTT CGG TTT AAA						387
Arg	His	Thr	Arg	Gly	Phe	Val
		30			35	40
AGC ATT GCG TTT TCT ATT TTG ATC ACT CTG TTG GTT ATT TTA GTG GAT						435
Ser	Ile	Ala	Phe	Ser	Ile	Leu
		45			50	55
ATT TGG GTT TAT AGC GAT GTG CGT CAG TTT TTA TTG GAC ACT TCT AGC						483
Ile	Trp	Val	Tyr	Ser	Asp	Val
				Arg	Gln	Phe
				Leu	Leu	Asp
				Thr	Ser	Ser

-217-

60	-65	70	
TCT TTT ATT TGG CTT TTG ATC GCT TTA CTA ATC AAG TGG GGC GTG ATT			531
Ser Phe Ile Trp Leu Leu Ile Ala Leu Leu Ile Lys Trp Gly Val Ile			
75	80	85	90
GTC ATA AGC GCA CGT AAA TGC TAC CAA TTC AGC CAA AAA ATG TTT ACG			579
Val Ile Ser Ala Arg Lys Cys Tyr Gln Phe Ser Gln Lys Met Phe Thr			
	95	100	105
CTC ATT CAA AGA AAA AGG CAA ATC AGA GAG AAT TTA AAA AAC CGC TCC			627
Leu Ile Gln Arg Lys Arg Gln Ile Arg Glu Asn Leu Lys Asn Arg Ser			
	110	115	120
AAC TAC AAA GAT ACC AAA AAT GCG GAA AAA CTC TCT AGC ATC GCT GAA			675
Asn Tyr Lys Asp Thr Lys Asn Ala Glu Lys Leu Ser Ser Ile Ala Glu			
	125	130	135
GAA ATC ATT TCA AAA AAA CAA GAA GAG TCC CGC CCC AAA GAA GAT TCT			723
Glu Ile Ile Ser Lys Lys Gln Glu Glu Ser Arg Pro Lys Glu Asp Ser			
	140	145	150
AAT CAT GAA AAC CAT AAA GAA AAG CTT TCT AAC ATT ACC GAA GAA AGT			771
Asn His Glu Asn His Lys Glu Lys Leu Ser Asn Ile Thr Glu Glu Ser			
	155	160	165
GAT TCT TAAAAAACAA GAGGAATTGA AAAGCTAAAA AGGATAGGGG GGGATTACCC AA			829
Asp Ser			
AGCATATTGG AGGG			843

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Met	Ser	Lys	Lys	Asn	Ser	Val	Ile	Ser	Gly	Leu	Met	Asn	Phe	Phe	Ser
1				5					10					15	
Glu	Lys	Asn	Glu	Arg	Trp	Leu	Leu	Ala	His	Arg	His	Thr	Arg	Gly	Phe
		20						25					30		
Val	Ile	Val	Ala	Trp	Leu	Phe	Arg	Phe	Lys	Ser	Ile	Ala	Phe	Ser	Ile
		35					40					45			
Leu	Ile	Thr	Leu	Leu	Val	Ile	Leu	Val	Asp	Ile	Trp	Val	Tyr	Ser	Asp
		50					55				60				

-218-

```

Val Arg Gln Phe Leu Leu- Asp Thr Ser Ser Ser Phe Ile Trp Leu Leu
65          70          75          80
Ile Ala Leu Leu Ile Lys Trp Gly Val Ile Val Ile Ser Ala Arg Lys
      85          90          95
Cys Tyr Gln Phe Ser Gln Lys Met Phe Thr Leu Ile Gln Arg Lys Arg
      100        105        110
Gln Ile Arg Glu Asn Leu Lys Asn Arg Ser Asn Tyr Lys Asp Thr Lys
      115        120        125
Asn Ala Glu Lys Leu Ser Ser Ile Ala Glu Glu Ile Ile Ser Lys Lys
      130        135        140
Gln Glu Glu Ser Arg Pro Lys Glu Asp Ser Asn His Glu Asn His Lys
145          150          155          160
Glu Lys Leu Ser Asn Ile Thr Glu Glu Ser Asp Ser
      165          170

```

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1047 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 34...1005
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

```

AGAAAGAAAC CATTCAAGGA ACGCATTGAT TTG ATG AAT AAA CCA TTT TTA ATC      54
                Met Asn Lys Pro Phe Leu Ile
                1                5

TTA CTC ATA GCC CTA ATT GTC TTT AGC GGC TGT AAC ATG AGA AAA TAT      102
Leu Leu Ile Ala Leu Ile Val Phe Ser Gly Cys Asn Met Arg Lys Tyr
      10                15                20

TTC AAA CCC GCT AAA CAC CAA ATT AAA GGC GAA GCG TAT TTC CCT AAC      150
Phe Lys Pro Ala Lys His Gln Ile Lys Gly Glu Ala Tyr Phe Pro Asn
      25                30                35

CAT TTG CAA GAA AGT ATC GTT TCG TCT AAT CGT TAT GGA GCC ATT TTG      198
His Leu Gln Glu Ser Ile Val Ser Ser Asn Arg Tyr Gly Ala Ile Leu
      40                45                50                55

AAA AAT GGA GCG GTT ATA GGC GAT AAA GGT TTA ACG CAG CTA AGA ATC      246
Lys Asn Gly Ala Val Ile Gly Asp Lys Gly Leu Thr Gln Leu Arg Ile
      60                65                70

GGT AAG AAC TTC AAT TAC GAA AGC AGT TTT TTA AAT GAG AGT CAA GGG      294

```

-219-

Gly	Lys	Asn	Phe	Asn	Tyr	Glu	Ser	Ser	Phe	Leu	Asn	Glu	Ser	Gln	Gly	
			75					80					85			
TTT	TTT	ATT	CTT	GCG	CAA	GAT	TGT	TTG	AAC	AAG	ATT	GAT	AAA	AAA	ACA	342
Phe	Phe	Ile	Leu	Ala	Gln	Asp	Cys	Leu	Asn	Lys	Ile	Asp	Lys	Lys	Thr	
		90					95					100				
AAC	AAA	AGC	AAG	GTG	GCT	AAG	ACT	GAA	GAA	ACG	GAA	TTG	AAA	TTA	AAG	390
Asn	Lys	Ser	Lys	Val	Ala	Lys	Thr	Glu	Glu	Thr		Glu	Leu	Lys	Leu	
	105					110						115				
GGC	GTT	GAA	GCG	GAA	GTC	CAA	GAT	AAA	GTC	TGT	CAT	CAA	GTG	GAA	TTG	438
Gly	Val	Glu	Ala	Glu	Val	Gln	Asp	Lys	Val	Cys	His	Gln	Val	Glu	Leu	
120					125					130					135	
ATT	AGC	AAT	AAC	CCT	AAC	GCC	AGC	CAA	CAA	TCT	ATC	GTT	ATT	CCT	TTG	486
Ile	Ser	Asn	Asn	Pro	Asn	Ala	Ser	Gln	Gln	Ser	Ile	Val	Ile	Pro	Leu	
				140					145					150		
GAG	ACT	TTT	GCC	TTG	AGC	GCA	AGC	GTT	AAA	GGG	AAT	CTT	TTA	GCG	GTG	534
Glu	Thr	Phe	Ala	Leu	Ser	Ala	Ser	Val	Lys	Gly	Asn	Leu	Leu	Ala	Val	
			155					160					165			
GTG	TTA	GCG	GAC	AAT	TCA	GCG	AAC	TTA	TAC	GAC	ATC	ACT	TCT	CAA	AAA	582
Val	Leu	Ala	Asp	Asn	Ser	Ala	Asn	Leu	Tyr	Asp	Ile	Thr	Ser	Gln	Lys	
		170					175					180				
TTG	CTT	TTT	AGT	GAG	AAA	GGT	TCC	CCA	AGC	ACC	ACG	ATC	AAT	TCT	TTA	630
Leu	Leu	Phe	Ser	Glu	Lys	Gly	Ser	Pro	Ser	Thr	Thr	Ile	Asn	Ser	Leu	
	185					190					195					
ATG	GCG	ATG	CCT	ATT	TTT	ATG	GAT	ACG	GTC	GTG	GTG	TTC	CCC	ATG	CTA	678
Met	Ala	Met	Pro	Ile	Phe	Met	Asp	Thr	Val	Val	Val	Phe	Pro	Met	Leu	
200					205					210				215		
GAT	GGG	CGC	TTG	TTG	GTC	GTG	GAT	TAT	GTG	CAC	GGA	AAC	CCT	ACG	CCT	726
Asp	Gly	Arg	Leu	Leu	Val	Val	Asp	Tyr	Val	His	Gly	Asn	Pro	Thr	Pro	
				220					225					230		
ATT	AGA	AAC	ATT	GTT	ATC	AGC	AGC	GAT	AAG	TTT	TTT	AAC	AAT	ATC	ACC	774
Ile	Arg	Asn	Ile	Val	Ile	Ser	Ser	Asp	Lys	Phe	Phe	Asn	Asn	Ile	Thr	
			235					240					245			
TAC	CTT	ATC	GTA	GAT	GGC	AAT	AAC	ATG	ATC	GCT	TCT	ACA	GGG	AAA	AGG	822
Tyr	Leu	Ile	Val	Asp	Gly	Asn	Asn	Met	Ile	Ala	Ser	Thr	Gly	Lys	Arg	
		250				255						260				
ATA	CTC	TCA	GTA	GTG	AGC	GGT	CAA	GAG	TTC	AAC	TAT	GAT	GGG	GAT	ATT	870
Ile	Leu	Ser	Val	Val	Ser	Gly	Gln	Glu	Phe	Asn	Tyr	Asp	Gly	Asp	Ile	
	265					270					275					
GTG	GAT	TTG	CTT	TAT	GAT	AAG	GGG	ACT	TTA	TAT	GTG	CTC	ACG	CTA	GAC	918
Val	Asp	Leu	Leu	Tyr	Asp	Lys	Gly	Thr	Leu	Tyr	Val	Leu	Thr	Leu	Asp	
280					285					290					295	

-220-

GGG CAG ATT TTG CAA ATG-GAT AAG AGT TTG AGG GAA TTA AAC AGC GTG 966
 Gly Gln Ile Leu Gln Met Asp Lys Ser Leu Arg Glu Leu Asn Ser Val
 300 305 310

AAA CTG CCT NTC NTC GCT CAA CAC GAT TGT ATT AAA CCA TAATAAATTG TA 1017
 Lys Leu Pro Xaa Xaa Ala Gln His Asp Cys Ile Lys Pro
 315 320

TTCTTTAGAA AAACGAGGGT ATGTGATAGA 1047

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Met	Asn	Lys	Pro	Phe	Leu	Ile	Leu	Leu	Ile	Ala	Leu	Ile	Val	Phe	Ser	1	5	10	15
Gly	Cys	Asn	Met	Arg	Lys	Tyr	Phe	Lys	Pro	Ala	Lys	His	Gln	Ile	Lys	20	25	30	
Gly	Glu	Ala	Tyr	Phe	Pro	Asn	His	Leu	Gln	Glu	Ser	Ile	Val	Ser	Ser	35	40	45	
Asn	Arg	Tyr	Gly	Ala	Ile	Leu	Lys	Asn	Gly	Ala	Val	Ile	Gly	Asp	Lys	50	55	60	
Gly	Leu	Thr	Gln	Leu	Arg	Ile	Gly	Lys	Asn	Phe	Asn	Tyr	Glu	Ser	Ser	65	70	75	80
Phe	Leu	Asn	Glu	Ser	Gln	Gly	Phe	Phe	Ile	Leu	Ala	Gln	Asp	Cys	Leu	85	90	95	
Asn	Lys	Ile	Asp	Lys	Lys	Thr	Asn	Lys	Ser	Lys	Val	Ala	Lys	Thr	Glu	100	105	110	
Glu	Thr	Glu	Leu	Lys	Leu	Lys	Gly	Val	Glu	Ala	Glu	Val	Gln	Asp	Lys	115	120	125	
Val	Cys	His	Gln	Val	Glu	Leu	Ile	Ser	Asn	Asn	Pro	Asn	Ala	Ser	Gln	130	135	140	
Gln	Ser	Ile	Val	Ile	Pro	Leu	Glu	Thr	Phe	Ala	Leu	Ser	Ala	Ser	Val	145	150	155	160
Lys	Gly	Asn	Leu	Leu	Ala	Val	Val	Leu	Ala	Asp	Asn	Ser	Ala	Asn	Leu	165	170	175	
Tyr	Asp	Ile	Thr	Ser	Gln	Lys	Leu	Leu	Phe	Ser	Glu	Lys	Gly	Ser	Pro	180	185	190	
Ser	Thr	Thr	Ile	Asn	Ser	Leu	Met	Ala	Met	Pro	Ile	Phe	Met	Asp	Thr	195	200	205	
Val	Val	Val	Phe	Pro	Met	Leu	Asp	Gly	Arg	Leu	Leu	Val	Val	Asp	Tyr	210	215	220	
Val	His	Gly	Asn	Pro	Thr	Pro	Ile	Arg	Asn	Ile	Val	Ile	Ser	Ser	Asp	225	230	235	240
Lys	Phe	Phe	Asn	Asn	Ile	Thr	Tyr	Leu	Ile	Val	Asp	Gly	Asn	Asn	Met	245	250	255	
Ile	Ala	Ser	Thr	Gly	Lys	Arg	Ile	Leu	Ser	Val	Val	Ser	Gly	Gln	Glu				

-221-

	260		265		270										
Phe	Asn	Tyr	Asp	Gly	Asp	Ile	Val	Asp	Leu	Leu	Tyr	Asp	Lys	Gly	Thr
	275						280					285			
Leu	Tyr	Val	Leu	Thr	Leu	Asp	Gly	Gln	Ile	Leu	Gln	Met	Asp	Lys	Ser
	290					295					300				
Leu	Arg	Glu	Leu	Asn	Ser	Val	Lys	Leu	Pro	Xaa	Xaa	Ala	Gln	His	Asp
305					310				315					320	
Cys	Ile	Lys	Pro												

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1968 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 153...1793
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 153...219
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

TCTGGGGGCA	TTGCTTACCC	TACTACTCGC	TTGAAACGCC	CAAGCCTGAT	CCAATCTCAT	60										
AAAGATTCTA	ATCGCAATTT	TAAAACCATC	ACTTTTGGC	TCGTTCCAC	AAAAAGCCAC	120										
GCAACTTACT	ACATCATTAA	GGTTTAATCA	CA	ATG	GAT	AAA	AAC	AAC	AAT	AAT	173					
				Met	Asp	Lys	Asn	Asn	Asn	Asn						
											-20					
CTC	CGC	TTG	ATT	TTA	GCG	ATC	GCT	CTG	TCT	TTC	TTG	TTT	ATC	GCT	CTT	221
Leu	Arg	Leu	Ile	Leu	Ala	Ile	Ala	Leu	Ser	Phe	Leu	Phe	Ile	Ala	Leu	
-15					-10					-5					1	
TAT	AGC	TAT	TTT	TTC	CAA	AAA	CCA	AAC	AAA	ACA	ACA	ACC	CAA	ACC	ACA	269
Tyr	Ser	Tyr	Phe	Phe	Gln	Lys	Pro	Asn	Lys	Thr	Thr	Thr	Gln	Thr	Thr	
			5					10					15			
AAG	CAA	GAA	ACA	ACC	AAC	AAC	CAT	ACA	GCA	ACA	AGT	CCT	AAC	GCG	CCC	317
Lys	Gln	Glu	Thr	Thr	Asn	Asn	His	Thr	Ala	Thr	Ser	Pro	Asn	Ala	Pro	
			20				25					30				
AAC	GCC	CAA	CAT	TTT	AGC	ACC	ACT	CAA	ACA	ACC	CCC	CAA	GAG	AAT	TTG	365
Asn	Ala	Gln	His	Phe	Ser	Thr	Thr	Gln	Thr	Thr	Pro	Gln	Glu	Asn	Leu	
			35				40				45					

-222-

CTA AGC ACG ATT TCT TTT GAG CAT GCC AGG ATT GAA ATT GAT TCT TTA	413
Leu Ser Thr Ile Ser Phe Glu His Ala Arg Ile Glu Ile Asp Ser Leu	
50 55 60 65	
GGG CGC ATC AAA CAG GTT TAT CTC AAG GAT AAA AAG TAT CTA ACC CCT	461
Gly Arg Ile Lys Gln Val Tyr Leu Lys Asp Lys Lys Tyr Leu Thr Pro	
70 75 80	
AAA CAA AAG GGC TTT TTA GAG CAT GTG GGC CAT CTT TTT AGC TCC AAA	509
Lys Gln Lys Gly Phe Leu Glu His Val Gly His Leu Phe Ser Ser Lys	
85 90 95	
GAA AAC GCG CAA CCC CCC CTA AAA GAG CTC CCC CTT TTA GCA GCC GAT	557
Glu Asn Ala Gln Pro Pro Leu Lys Glu Leu Pro Leu Leu Ala Ala Asp	
100 105 110	
AAA CTC AAG CCT TTA GAA GTG CGT TTT TTA GAC CCT ACG CTC AAT AAC	605
Lys Leu Lys Pro Leu Glu Val Arg Phe Leu Asp Pro Thr Leu Asn Asn	
115 120 125	
AAA GCG TTC AAC ACC CCT TAT AGC GCT TCA AAA ACC ACT CTT GGG CCT	653
Lys Ala Phe Asn Thr Pro Tyr Ser Ala Ser Lys Thr Thr Leu Gly Pro	
130 135 140 145	
AAC GAA CAG CTT GTT TTA ACC CAA GAT TTA GGC ACT CTT AGC ATC ATT	701
Asn Glu Gln Leu Val Leu Thr Gln Asp Leu Gly Thr Leu Ser Ile Ile	
150 155 160	
AAA ACC CTG ACT TTC TAT GAT GAT TTG CAT TAT GAT TTA AAA ATC GCA	749
Lys Thr Leu Thr Phe Tyr Asp Asp Leu His Tyr Asp Leu Lys Ile Ala	
165 170 175	
TTC AAA TCG CCC AAT AAC CTT ATC CCT AGC TAT GTG ATC ACC AAT GGT	797
Phe Lys Ser Pro Asn Asn Leu Ile Pro Ser Tyr Val Ile Thr Asn Gly	
180 185 190	
TAC AGG CCG GTG GCT GAT TTG GAC AGC TAC ACC TTT TCA GGC GTG CTT	845
Tyr Arg Pro Val Ala Asp Leu Asp Ser Tyr Thr Phe Ser Gly Val Leu	
195 200 205	
TTA GAA AAT AGC GAC AAA AAA ATT GAA AAA ATT GAA GAT AAA GAC GCT	893
Leu Glu Asn Ser Asp Lys Lys Ile Glu Lys Ile Glu Asp Lys Asp Ala	
210 215 220 225	
AAA GAA ATC AAA CGC TTT TCT AAC ACC CTC TTT TTA TCC AGC GTG GAT	941
Lys Glu Ile Lys Arg Phe Ser Asn Thr Leu Phe Leu Ser Ser Val Asp	
230 235 240	
AGG TAT TTC ACC ACC TTG CTT TTC ACT AAA GAT CCT CAA GGT TTT GAA	989
Arg Tyr Phe Thr Thr Leu Leu Phe Thr Lys Asp Pro Gln Gly Phe Glu	
245 250 255	
GCC TTA ATT GAT TCA GAA ATC GGC ACT AAA AAC CCC TTA GGG TTC ATT	1037
Ala Leu Ile Asp Ser Glu Ile Gly Thr Lys Asn Pro Leu Gly Phe Ile	
260 265 270	

-223-

TCC CTT AAA AAT GAA GCG -AAT TTG CAT GGC TAT ATT GGC CCT AAG GAT	1085
Ser Leu Lys Asn Glu Ala Asn Leu His Gly Tyr Ile Gly Pro Lys Asp	
275 280 285	
TAC CGC TCT TTG AAA GCG ATT TCA CCC ATG CTC ACC GAT GTG ATA GAG	1133
Tyr Arg Ser Leu Lys Ala Ile Ser Pro Met Leu Thr Asp Val Ile Glu	
290 295 300 305	
TAT GGC TTA ATC ACT TTC TTT GCA AAA GGC GTG TTT GTT TTA CTG GAT	1181
Tyr Gly Leu Ile Thr Phe Phe Ala Lys Gly Val Phe Val Leu Leu Asp	
310 315 320	
TAT TTG TAT CAA TTC GTG GGC AAT TGG GGT TGG GCT ATC ATT CTT TTA	1229
Tyr Leu Tyr Gln Phe Val Gly Asn Trp Gly Trp Ala Ile Ile Leu Leu	
325 330 335	
ACG ATT ATC GTG CGC ATC ATC CTT TAT CCT TTA AGC TAT AAG GGC ATG	1277
Thr Ile Ile Val Arg Ile Ile Leu Tyr Pro Leu Ser Tyr Lys Gly Met	
340 345 350	
GTG AGC ATG CAA AAG CTC AAA GAA TTA GCC CCT AAA ATG AAA GAA CTC	1325
Val Ser Met Gln Lys Leu Lys Glu Leu Ala Pro Lys Met Lys Glu Leu	
355 360 365	
CAA GAA AAA TAC AAG GGC GAA CCC CAA AAA TTG CAA GCC CAC ATG ATG	1373
Gln Glu Lys Tyr Lys Gly Glu Pro Gln Lys Leu Gln Ala His Met Met	
370 375 380 385	
CAG CTT TAC AAA AAA CAT GGG GCT AAC CCA CTA GGG GGT TGT CTG CCC	1421
Gln Leu Tyr Lys Lys His Gly Ala Asn Pro Leu Gly Gly Cys Leu Pro	
390 395 400	
TTA ATC TTA CAA ATC CCG GTG TTT TTT GCC ATT TAT AGA GTG CTT TAT	1469
Leu Ile Leu Gln Ile Pro Val Phe Phe Ala Ile Tyr Arg Val Leu Tyr	
405 410 415	
AAC GCT GTG GAA TTG AAA AGC TCA GAG TGG ATC TTA TGG ATT CAT GAT	1517
Asn Ala Val Glu Leu Lys Ser Ser Glu Trp Ile Leu Trp Ile His Asp	
420 425 430	
TTA TCC ATC ATG GAT CCG TAT TTT ATT TTA CCG CTT CTT ATG GGA GCG	1565
Leu Ser Ile Met Asp Pro Tyr Phe Ile Leu Pro Leu Leu Met Gly Ala	
435 440 445	
TCT ATG TAT TGG CAC CAA AGC GTT ACG CCA AAC ACC ATG ACC GAT CCC	1613
Ser Met Tyr Trp His Gln Ser Val Thr Pro Asn Thr Met Thr Asp Pro	
450 455 460 465	
ATG CAA GCA AAG ATT TTT AAA CTC TTA CCC CTA TTA TTC ACA ATC TTT	1661
Met Gln Ala Lys Ile Phe Lys Leu Leu Pro Leu Leu Phe Thr Ile Phe	
470 475 480	
TTA ATC ACT TTC CCG GCA GGG TTA GTC TTG TAT TGG ACC ACG AAC AAC	1709
Leu Ile Thr Phe Pro Ala Gly Leu Val Leu Tyr Trp Thr Thr Asn Asn	
485 490 495	

-224-

```

ATC CTT TCG GTG TTG CAA CAA CTC ATC ATC AAT AAA GTC TTA GAG AAT      1757
Ile Leu Ser Val Leu Gln Gln Leu Ile Ile Asn Lys Val Leu Glu Asn
      500                      505                      510

AAA AAA CGC ATG CAT GCG CAA AAC AAA AAG GAA CAT TGATGCAAAA TTTTAT      1809
Lys Lys Arg Met His Ala Gln Asn Lys Lys Glu His
      515                      520                      525

TGAAATCAAA GCCAAAACCT TAGAAGAAGC CCTCATTCAA GCTTCTATCG CCTTGAATTG      1869
CCCCATTATT AATTTGCAAT ACGAAGTCAT TCAAACGCCC TCTAAAGGGT TTTTAAGCAT      1929
TGGTAAAAAA GAAGCCATTA TCTTAGCGGG CGTTAAAGA                          1968

```

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 547 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

(A) NAME/KEY: Signal Sequence

(B) LOCATION: 1...22

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```

Met Asp Lys Asn Asn Asn Asn Leu Arg Leu Ile Leu Ala Ile Ala Leu
      -20                      -15                      -10
Ser Phe Leu Phe Ile Ala Leu Tyr Ser Tyr Phe Phe Gln Lys Pro Asn
      -5                      1                      5                      10
Lys Thr Thr Thr Gln Thr Thr Lys Gln Glu Thr Thr Asn Asn His Thr
      15                      20                      25
Ala Thr Ser Pro Asn Ala Pro Asn Ala Gln His Phe Ser Thr Thr Gln
      30                      35                      40
Thr Thr Pro Gln Glu Asn Leu Leu Ser Thr Ile Ser Phe Glu His Ala
      45                      50                      55
Arg Ile Glu Ile Asp Ser Leu Gly Arg Ile Lys Gln Val Tyr Leu Lys
      60                      65                      70
Asp Lys Lys Tyr Leu Thr Pro Lys Gln Lys Gly Phe Leu Glu His Val
      75                      80                      85                      90
Gly His Leu Phe Ser Ser Lys Glu Asn Ala Gln Pro Pro Leu Lys Glu
      95                      100                      105
Leu Pro Leu Leu Ala Ala Asp Lys Leu Lys Pro Leu Glu Val Arg Phe
      110                      115                      120
Leu Asp Pro Thr Leu Asn Asn Lys Ala Phe Asn Thr Pro Tyr Ser Ala
      125                      130                      135
Ser Lys Thr Thr Leu Gly Pro Asn Glu Gln Leu Val Leu Thr Gln Asp
      140                      145                      150
Leu Gly Thr Leu Ser Ile Ile Lys Thr Leu Thr Phe Tyr Asp Asp Leu

```

-225-

155					160					165				170
His	Tyr	Asp	Leu	Lys	Ile	Ala	Phe	Lys	Ser	Pro	Asn	Asn	Leu	Ile
				175					180					185
Ser	Tyr	Val	Ile	Thr	Asn	Gly	Tyr	Arg	Pro	Val	Ala	Asp	Leu	Asp
			190					195					200	
Tyr	Thr	Phe	Ser	Gly	Val	Leu	Leu	Glu	Asn	Ser	Asp	Lys	Lys	Ile
		205					210					215		
Lys	Ile	Glu	Asp	Lys	Asp	Ala	Lys	Glu	Ile	Lys	Arg	Phe	Ser	Asn
	220					225					230			
Leu	Phe	Leu	Ser	Ser	Val	Asp	Arg	Tyr	Phe	Thr	Thr	Leu	Leu	Phe
235					240					245				250
Lys	Asp	Pro	Gln	Gly	Phe	Glu	Ala	Leu	Ile	Asp	Ser	Glu	Ile	Gly
			255						260					265
Lys	Asn	Pro	Leu	Gly	Phe	Ile	Ser	Leu	Lys	Asn	Glu	Ala	Asn	Leu
		270						275					280	His
Gly	Tyr	Ile	Gly	Pro	Lys	Asp	Tyr	Arg	Ser	Leu	Lys	Ala	Ile	Ser
	285						290					295		Pro
Met	Leu	Thr	Asp	Val	Ile	Glu	Tyr	Gly	Leu	Ile	Thr	Phe	Phe	Ala
	300					305					310			Lys
Gly	Val	Phe	Val	Leu	Leu	Asp	Tyr	Leu	Tyr	Gln	Phe	Val	Gly	Asn
315					320					325				330
Gly	Trp	Ala	Ile	Ile	Leu	Leu	Thr	Ile	Ile	Val	Arg	Ile	Ile	Leu
			335						340					345
Pro	Leu	Ser	Tyr	Lys	Gly	Met	Val	Ser	Met	Gln	Lys	Leu	Lys	Glu
		350						355					360	Leu
Ala	Pro	Lys	Met	Lys	Glu	Leu	Gln	Glu	Lys	Tyr	Lys	Gly	Glu	Pro
	365						370					375		Gln
Lys	Leu	Gln	Ala	His	Met	Met	Gln	Leu	Tyr	Lys	Lys	His	Gly	Ala
	380				385						390			Asn
Pro	Leu	Gly	Gly	Cys	Leu	Pro	Leu	Ile	Leu	Gln	Ile	Pro	Val	Phe
395				400					405					410
Ala	Ile	Tyr	Arg	Val	Leu	Tyr	Asn	Ala	Val	Glu	Leu	Lys	Ser	Ser
			415					420					425	Glu
Trp	Ile	Leu	Trp	Ile	His	Asp	Leu	Ser	Ile	Met	Asp	Pro	Tyr	Phe
	430						435					440		Ile
Leu	Pro	Leu	Leu	Met	Gly	Ala	Ser	Met	Tyr	Trp	His	Gln	Ser	Val
	445						450					455		Thr
Pro	Asn	Thr	Met	Thr	Asp	Pro	Met	Gln	Ala	Lys	Ile	Phe	Lys	Leu
	460					465				470				Leu
Pro	Leu	Leu	Phe	Thr	Ile	Phe	Leu	Ile	Thr	Phe	Pro	Ala	Gly	Leu
475				480					485					490
Leu	Tyr	Trp	Thr	Thr	Asn	Asn	Ile	Leu	Ser	Val	Leu	Gln	Gln	Leu
			495					500					505	Ile
Ile	Asn	Lys	Val	Leu	Glu	Asn	Lys	Lys	Arg	Met	His	Ala	Gln	Asn
			510				515						520	Lys
Lys	Glu	His												
		525												

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3280 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

-226-

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 151...3207
 (D) OTHER INFORMATION:

(A) NAME/KEY: Signal Sequence
 (B) LOCATION: 151...241
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

TAAAGGTTTT	AGGCCTGTGG	TGGTTCAAGT	TTAGAAGAG	CGCAGCAAGA	TTTTTATCGT	60
GAACGCTCAA	AATTTACACC	CTAATGACAG	CGTGGCAGTG	GGGTCATTGA	TAGGGTTAAA	120
AGGCATGATC	AACAATTTAG	GGGAGGAATG	ATG CTC GCT TCC ATT ATT GAA TTT			174
			Met Leu Ala Ser Ile Ile Glu Phe			
			-30	-25		
TCC TTA CGC CAA AGA GTG ATC GTG ATT GTT GGT GCG ATT CTT ATT TTA						222
Ser Leu Arg Gln Arg Val Ile Val Ile Val Gly Ala Ile Leu Ile Leu						
	-20		-15		-10	
TTT TTT GGG ACT TAT AGT TTT ATC AAC ACT CCA GTG GAC GCT TTC CCG						270
Phe Phe Gly Thr Tyr Ser Phe Ile Asn Thr Pro Val Asp Ala Phe Pro						
	-5		1		5	10
GAT ATT TCG CCC ACT CAA GTT AAA ATC ATT TTA AAA CTC CCC GGC TCT						318
Asp Ile Ser Pro Thr Gln Val Lys Ile Ile Leu Lys Leu Pro Gly Ser						
		15		20	25	
AGC CCT GAA GAA ATG GAA AAC AAC ATC GTG CGC CCT TTA GAA TTG GAG						366
Ser Pro Glu Glu Met Glu Asn Asn Ile Val Arg Pro Leu Glu Leu Glu						
	30		35		40	
CTT TTA GGC TTG AAA GGG CAA AAA TCT TTA AGG AGT GTT TCA AAA TAT						414
Leu Leu Gly Leu Lys Gly Gln Lys Ser Leu Arg Ser Val Ser Lys Tyr						
	45		50		55	
TCT ATT TCA GAT ATT ACG ATA GAT TTT GAT GAC AGC GTG GAT ATT TAT						462
Ser Ile Ser Asp Ile Thr Ile Asp Phe Asp Asp Ser Val Asp Ile Tyr						
	60		65		70	
TTA GCG AGG AAT ATT GTC AAT GAG CGC TTG AGC AGC GTG ATG AAA GAT						510
Leu Ala Arg Asn Ile Val Asn Glu Arg Leu Ser Ser Val Met Lys Asp						
	75		80		85	90
TTA CCC GTG GGG GTT GAG GGG GGC ATG GCG CCC ATT GTT ACG CCG CTA						558
Leu Pro Val Gly Val Glu Gly Gly Met Ala Pro Ile Val Thr Pro Leu						
	95		100		105	

-227-

TCA GAT ATC TTT ATG TTC ACT ATT GAT GGC AAT ATC ACT GAG ATA GAA Ser Asp Ile Phe Met Phe Thr Ile Asp Gly Asn Ile Thr Glu Ile Glu 110 115 120	606
AAA CGA CAG CTT TTA GAT TTT GTG ATC CGC CCA CAA TTA AGA ATG ATT Lys Arg Gln Leu Leu Asp Phe Val Ile Arg Pro Gln Leu Arg Met Ile 125 130 135	654
AGC GGC GTA GCA GAT GTC AAT TCC ATT GGA GGC TTT AGC AGA GCG TTT Ser Gly Val Ala Asp Val Asn Ser Ile Gly Gly Phe Ser Arg Ala Phe 140 145 150	702
GTG ATC GTG CCG GAT TTT AAT GAC ATG GCA AGG CTT GGG GTG AGT ATT Val Ile Val Pro Asp Phe Asn Asp Met Ala Arg Leu Gly Val Ser Ile 155 160 165 170	750
TCT GAT TTA GAA TCG GCT GTG AGA GTG AAT TTA AGA AAC AGC GGA GCG Ser Asp Leu Glu Ser Ala Val Arg Val Asn Leu Arg Asn Ser Gly Ala 175 180 185	798
GGG CGC GTG GAT AGA GAT GGC GAA ACC TTT TTA GTC AAA ATC CAA ACC Gly Arg Val Asp Arg Asp Gly Glu Thr Phe Leu Val Lys Ile Gln Thr 190 195 200	846
GCT TCT TTG AGT TTA GAA GAC ATT GGC AAA ATC ACC GTT TCC ACT AAT Ala Ser Leu Ser Leu Glu Asp Ile Gly Lys Ile Thr Val Ser Thr Asn 205 210 215	894
TTA GGG CAT TTG CAC ATT AAG GAT TTT GCG AAA GTC ATC AGC CAG TCT Leu Gly His Leu His Ile Lys Asp Phe Ala Lys Val Ile Ser Gln Ser 220 225 230	942
CGC ACC CGT TTG GGG TTT GTT ACT AAA GAT GGC GTG GGC GAG ACC ACA Arg Thr Arg Leu Gly Phe Val Thr Lys Asp Gly Val Gly Glu Thr Thr 235 240 245 250	990
GAA GGC TTG GTG CTT TCT TTA AAA GAC GCT AAC ACC AAA GAA ATC ATC Glu Gly Leu Val Leu Ser Leu Lys Asp Ala Asn Thr Lys Glu Ile Ile 255 260 265	1038
ACT CAA GTG TAT CAA AAA CTA GAA GAA TTA AAA CCC TTT TTA CCG AAT Thr Gln Val Tyr Gln Lys Leu Glu Glu Leu Lys Pro Phe Leu Pro Asn 270 275 280	1086
GGC GTG TCC ATT AAT GTT TTT TAT GAT CGC TCA GAA TTT ACG CAA AAA Gly Val Ser Ile Asn Val Phe Tyr Asp Arg Ser Glu Phe Thr Gln Lys 285 290 295	1134
GCC ATT GCC ACC GTT TCT AAA ACC CTC ATT GAA GCC GTT GTT TTA ATC Ala Ile Ala Thr Val Ser Lys Thr Leu Ile Glu Ala Val Val Leu Ile 300 305 310	1182
ATC ATC ACG CTC TTT TTA TTT TTA GGG AAT TTG AGG GCG AGC GTG GCT Ile Ile Thr Leu Phe Leu Phe Leu Gly Asn Leu Arg Ala Ser Val Ala 315 320 325 330	1230

-228-

GTG GGG GTG ATT TTA CCT TTA AGC TTG TCC GTG GCG TTT ATT TTT ATC	1278
Val Gly Val Ile Leu Pro Leu Ser Leu Ser Val Ala Phe Ile Phe Ile	
335 340 345	
AAG TTT AGC GAT CTG ACT TTA AAT TTG ATG AGT TTA GGG GGA TTG GTT	1326
Lys Phe Ser Asp Leu Thr Leu Asn Leu Met Ser Leu Gly Gly Leu Val	
350 355 360	
ATC GCT ATA GGC ATG CTC ATT GAC TCA GCC GTG GTG GTG GTG GAA AAC	1374
Ile Ala Ile Gly Met Leu Ile Asp Ser Ala Val Val Val Val Glu Asn	
365 370 375	
GCT TTT GAA AAA TTA AGC GCT AAC ACT AAA ACC ACT AAA CTC CAT GCA	1422
Ala Phe Glu Lys Leu Ser Ala Asn Thr Lys Thr Thr Lys Leu His Ala	
380 385 390	
ATC TAT CGT TCG TGT AAA GAA ATC GCT GTT TCA GTG GTG AGC GGG GTG	1470
Ile Tyr Arg Ser Cys Lys Glu Ile Ala Val Ser Val Val Ser Gly Val	
395 400 405 410	
GTG ATC ATC ATT GTG TTT TTT GTG CCG ATT TTA ACC TTA CAG GGG TTA	1518
Val Ile Ile Ile Val Phe Phe Val Pro Ile Leu Thr Leu Gln Gly Leu	
415 420 425	
GAG GGT AAG ATG TTT AGG CCT TTA GCG CAA AGC ATT GTG TAT GCG CTT	1566
Glu Gly Lys Met Phe Arg Pro Leu Ala Gln Ser Ile Val Tyr Ala Leu	
430 435 440	
TTA GGC ACT TTA GTT CTA TCT ATT ACA ATC ATT CCT GTA GTC AGC TCT	1614
Leu Gly Thr Leu Val Leu Ser Ile Thr Ile Ile Pro Val Val Ser Ser	
445 450 455	
CTT GTC TTA AAA GCC ACG CCC CAT AGC GAA ACC TTT TTA ACG AGG TTT	1662
Leu Val Leu Lys Ala Thr Pro His Ser Glu Thr Phe Leu Thr Arg Phe	
460 465 470	
TTA AAC AGA ATC TAC GCC CCT TTA TTG GAA TTT TTT GTG CAT AAC CCT	1710
Leu Asn Arg Ile Tyr Ala Pro Leu Leu Glu Phe Phe Val His Asn Pro	
475 480 485 490	
AAA AAA GTG ATT TTA GGA GCG TTT GTT TTT TTA ATC GCA AGC CTT TCT	1758
Lys Lys Val Ile Leu Gly Ala Phe Val Phe Leu Ile Ala Ser Leu Ser	
495 500 505	
TTA TTC CCT TTT GTG GGG AAG AAT TTC ATG CCC GTT TTA GAT GAG GGC	1806
Leu Phe Pro Phe Val Gly Lys Asn Phe Met Pro Val Leu Asp Glu Gly	
510 515 520	
GAT GTG GTT TTG AGC GTG GAA ACC ACC CCT TCT ATT TCT TTA GAT CAA	1854
Asp Val Val Leu Ser Val Glu Thr Thr Pro Ser Ile Ser Leu Asp Gln	
525 530 535	
TCT AGG GAT CTC ATG CTA AAC ATT GAG AGC GCG ATT AAA AAG CAT GTC	1902
Ser Arg Asp Leu Met Leu Asn Ile Glu Ser Ala Ile Lys Lys His Val	
540 545 550	

-229-

AAG GAA GTT AAA AGC ATT GTC GCG CGC ACA GGG AGC GAT GAA TTG GGG	1950
Lys Glu Val Lys Ser Ile Val Ala Arg Thr Gly Ser Asp Glu Leu Gly	
555 560 565 570	
CTG GAT TTA GGA GGT TTG AAT CAA ACC GAT ACT TTT ATT TCT TTT ATT	1998
Leu Asp Leu Gly Gly Leu Asn Gln Thr Asp Thr Phe Ile Ser Phe Ile	
575 580 585	
CCT AAA AAA GAA TGG AGC GTT AAA ACC AAA GAT GAA TTA TTA GAA AAA	2046
Pro Lys Lys Glu Trp Ser Val Lys Thr Lys Asp Glu Leu Leu Glu Lys	
590 595 600	
ATC ATG GAT TCT TTA AAA GAC TTT AAG GGG ATT AAC TTT TCT TTC ACC	2094
Ile Met Asp Ser Leu Lys Asp Phe Lys Gly Ile Asn Phe Ser Phe Thr	
605 610 615	
CAA CCC ATT GAA ATG AGA ATT TCT GAA ATG CTG ACA GGG GTT AGG GGG	2142
Gln Pro Ile Glu Met Arg Ile Ser Glu Met Leu Thr Gly Val Arg Gly	
620 625 630	
GAT TTA GCG GTT AAG ATT TTT GGA GAT GGT ATT AGC GAA TTG AAT GAA	2190
Asp Leu Ala Val Lys Ile Phe Gly Asp Gly Ile Ser Glu Leu Asn Glu	
635 640 645 650	
TTG AGT TTT CAA ATC GCG CAA GCT CTA AAA GGG ATT AAA GGA TCT AGT	2238
Leu Ser Phe Gln Ile Ala Gln Ala Leu Lys Gly Ile Lys Gly Ser Ser	
655 660 665	
GAA GTT TTA ACC ACG CTT AAT GAG GGC GTG AAT TAT TTG TAT GTA ACC	2286
Glu Val Leu Thr Thr Leu Asn Glu Gly Val Asn Tyr Leu Tyr Val Thr	
670 675 680	
CCT AAT AAA GAA TCG ATG GCG GAT GTG GGG ATC ACT AGC GAT GAA TTT	2334
Pro Asn Lys Glu Ser Met Ala Asp Val Gly Ile Thr Ser Asp Glu Phe	
685 690 695	
TCC AAG TTT TTA AAA TCC GCT TTA GAG GGC TTG GTT GTA GAT GTG ATC	2382
Ser Lys Phe Leu Lys Ser Ala Leu Glu Gly Leu Val Val Asp Val Ile	
700 705 710	
CCT ACA GGG ATT TCA CGC ACG CCA GTG ATG ATC CGC CAA GAG AGC GAT	2430
Pro Thr Gly Ile Ser Arg Thr Pro Val Met Ile Arg Gln Glu Ser Asp	
715 720 725 730	
TTT GCA AGC TCT ATC ACT AAA ATC AAA AGT TTA GCC TTG ACT TCA AAA	2478
Phe Ala Ser Ser Ile Thr Lys Ile Lys Ser Leu Ala Leu Thr Ser Lys	
735 740 745	
TAT GGC GTT TTA GTG CCT ATC ACT TCT ATC GCC AAA ATT GAA GAA GTG	2526
Tyr Gly Val Leu Val Pro Ile Thr Ser Ile Ala Lys Ile Glu Glu Val	
750 755 760	
GAT GGC CCT GTT TCT GTT GTG CGT GAA AAT TCA ATG CGC ATG AGC GTG	2574
Asp Gly Pro Val Ser Val Val Arg Glu Asn Ser Met Arg Met Ser Val	
765 770 775	

-230-

GTT CGC AGT AAT GTG GTG GGG CGC GAT TTG AAA TCT TTT GTA GAA GAG	2622
Val Arg Ser Asn Val Val Gly Arg Asp Leu Lys Ser Phe Val Glu Glu	
780 785 790	
GCT AAA AAA GTG ATC GCT CAA AAC ATC AAA CTC CCT CCC AGC TAC TAT	2670
Ala Lys Lys Val Ile Ala Gln Asn Ile Lys Leu Pro Pro Ser Tyr Tyr	
795 800 805 810	
ATC ACT TAT GGG GGG CAG TTT GAA AAC CAG CAA CGG GCC AAT AAA AGG	2718
Ile Thr Tyr Gly Gly Gln Phe Glu Asn Gln Arg Ala Asn Lys Arg	
815 820 825	
CTC TCC ACC GTT ATC CCT TTA AGC ATC TTA GCG ATT TTT TTC ATT CTT	2766
Leu Ser Thr Val Ile Pro Leu Ser Ile Leu Ala Ile Phe Phe Ile Leu	
830 835 840	
TTT TTC ACT TTT AAA AGC ATT CCT TTA GCC TTG CTC ATT CTT TTG AAT	2814
Phe Phe Thr Phe Lys Ser Ile Pro Leu Ala Leu Leu Ile Leu Leu Asn	
845 850 855	
ATC CCT TTT GCG GTT ACC GGA GGC CTT ATT GCG TTG TTT GCG GTC GGG	2862
Ile Pro Phe Ala Val Thr Gly Gly Leu Ile Ala Leu Phe Ala Val Gly	
860 865 870	
GAG TAT ATT TCA GTG CCA GCG AGC GTG GGC TTT ATC GCT CTT TTT GGG	2910
Glu Tyr Ile Ser Val Pro Ala Ser Val Gly Phe Ile Ala Leu Phe Gly	
875 880 885 890	
ATT GCG GTT TTA AAT GGC GTG GTG ATG ATA GGC TAT TTT AAA GAG CTT	2958
Ile Ala Val Leu Asn Gly Val Val Met Ile Gly Tyr Phe Lys Glu Leu	
895 900 905	
CTC TTG CAA GGG AAA AGC GTA GAA GAA TGC GTT TTA TTG GGC GCT AAA	3006
Leu Leu Gln Gly Lys Ser Val Glu Glu Cys Val Leu Leu Gly Ala Lys	
910 915 920	
AGG CGT TTG AGA CCG GTT TTA ATG ACC GCT TGC ATT GCC GGT TTG GGT	3054
Arg Arg Leu Arg Pro Val Leu Met Thr Ala Cys Ile Ala Gly Leu Gly	
925 930 935	
TTG CTC CCT TTA TTA TTT TCT CAT AGC GTG GGA TCA GAA GTC CAA AAA	3102
Leu Leu Pro Leu Leu Phe Ser His Ser Val Gly Ser Glu Val Gln Lys	
940 945 950	
CCT TTA GCG ATC GTG GTG CTT GGA GGC TTG GTT ACC TCA AGC GCT CTA	3150
Pro Leu Ala Ile Val Val Leu Gly Gly Leu Val Thr Ser Ser Ala Leu	
955 960 965 970	
ACC TTA CTC CTA CTG CCG CCA ATG TTT ATG CTC ATC GCT AAA AAG ATT	3198
Thr Leu Leu Leu Leu Pro Pro Met Phe Met Leu Ile Ala Lys Lys Ile	
975 980 985	
AAA ATC GTT TGAGTTAAAG GATTTCACAT GCTCGCTTTA GAAATTTATA TTGATATTT	3256
Lys Ile Val	

-231-

GTTTGAAAGA CGCTTTAATA GATT

3280

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1019 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

(A) NAME/KEY: Signal Sequence

(B) LOCATION: 1...30

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

```

Met Leu Ala Ser Ile Ile Glu Phe Ser Leu Arg Gln Arg Val Ile Val
-30          -25          -20          -15
Ile Val Gly Ala Ile Leu Ile Leu Phe Phe Gly Thr Tyr Ser Phe Ile
          -10          -5          1
Asn Thr Pro Val Asp Ala Phe Pro Asp Ile Ser Pro Thr Gln Val Lys
      5          10          15
Ile Ile Leu Lys Leu Pro Gly Ser Ser Pro Glu Glu Met Glu Asn Asn
      20          25          30
Ile Val Arg Pro Leu Glu Leu Glu Leu Leu Gly Leu Lys Gly Gln Lys
      35          40          45          50
Ser Leu Arg Ser Val Ser Lys Tyr Ser Ile Ser Asp Ile Thr Ile Asp
          55          60          65
Phe Asp Asp Ser Val Asp Ile Tyr Leu Ala Arg Asn Ile Val Asn Glu
          70          75          80
Arg Leu Ser Ser Val Met Lys Asp Leu Pro Val Gly Val Glu Gly Gly
          85          90          95
Met Ala Pro Ile Val Thr Pro Leu Ser Asp Ile Phe Met Phe Thr Ile
          100          105          110
Asp Gly Asn Ile Thr Glu Ile Glu Lys Arg Gln Leu Leu Asp Phe Val
          115          120          125          130
Ile Arg Pro Gln Leu Arg Met Ile Ser Gly Val Ala Asp Val Asn Ser
          135          140          145
Ile Gly Gly Phe Ser Arg Ala Phe Val Ile Val Pro Asp Phe Asn Asp
          150          155          160
Met Ala Arg Leu Gly Val Ser Ile Ser Asp Leu Glu Ser Ala Val Arg
          165          170          175
Val Asn Leu Arg Asn Ser Gly Ala Gly Arg Val Asp Arg Asp Gly Glu
          180          185          190
Thr Phe Leu Val Lys Ile Gln Thr Ala Ser Leu Ser Leu Glu Asp Ile
          195          200          205          210
Gly Lys Ile Thr Val Ser Thr Asn Leu Gly His Leu His Ile Lys Asp
          215          220          225
Phe Ala Lys Val Ile Ser Gln Ser Arg Thr Arg Leu Gly Phe Val Thr

```


Lys	Asp	Gly	Val	Gly	Glu	Thr	Thr	Glu	Gly	Leu	Val	Leu	Ser	Leu	Lys
		245					250					255			
Asp	Ala	Asn	Thr	Lys	Glu	Ile	Ile	Thr	Gln	Val	Tyr	Gln	Lys	Leu	Glu
	260					265					270				
Glu	Leu	Lys	Pro	Phe	Leu	Pro	Asn	Gly	Val	Ser	Ile	Asn	Val	Phe	Tyr
275					280					285					290
Asp	Arg	Ser	Glu	Phe	Thr	Gln	Lys	Ala	Ile	Ala	Thr	Val	Ser	Lys	Thr
				295				300						305	
Leu	Ile	Glu	Ala	Val	Val	Leu	Ile	Ile	Ile	Thr	Leu	Phe	Leu	Phe	Leu
			310					315					320		
Gly	Asn	Leu	Arg	Ala	Ser	Val	Ala	Val	Gly	Val	Ile	Leu	Pro	Leu	Ser
		325					330					335			
Leu	Ser	Val	Ala	Phe	Ile	Phe	Ile	Lys	Phe	Ser	Asp	Leu	Thr	Leu	Asn
	340					345					350				
Leu	Met	Ser	Leu	Gly	Gly	Leu	Val	Ile	Ala	Ile	Gly	Met	Leu	Ile	Asp
355				360						365					370
Ser	Ala	Val	Val	Val	Val	Glu	Asn	Ala	Phe	Glu	Lys	Leu	Ser	Ala	Asn
				375					380					385	
Thr	Lys	Thr	Thr	Lys	Leu	His	Ala	Ile	Tyr	Arg	Ser	Cys	Lys	Glu	Ile
			390					395					400		
Ala	Val	Ser	Val	Val	Ser	Gly	Val	Val	Ile	Ile	Ile	Val	Phe	Phe	Val
		405					410					415			
Pro	Ile	Leu	Thr	Leu	Gln	Gly	Leu	Glu	Gly	Lys	Met	Phe	Arg	Pro	Leu
	420				425						430				
Ala	Gln	Ser	Ile	Val	Tyr	Ala	Leu	Leu	Gly	Thr	Leu	Val	Leu	Ser	Ile
435				440						445					450
Thr	Ile	Ile	Pro	Val	Val	Ser	Ser	Leu	Val	Leu	Lys	Ala	Thr	Pro	His
			455					460						465	
Ser	Glu	Thr	Phe	Leu	Thr	Arg	Phe	Leu	Asn	Arg	Ile	Tyr	Ala	Pro	Leu
			470					475				480			
Leu	Glu	Phe	Phe	Val	His	Asn	Pro	Lys	Lys	Val	Ile	Leu	Gly	Ala	Phe
		485					490					495			
Val	Phe	Leu	Ile	Ala	Ser	Leu	Ser	Leu	Phe	Pro	Phe	Val	Gly	Lys	Asn
	500				505						510				
Phe	Met	Pro	Val	Leu	Asp	Glu	Gly	Asp	Val	Val	Leu	Ser	Val	Glu	Thr
515				520						525					530
Thr	Pro	Ser	Ile	Ser	Leu	Asp	Gln	Ser	Arg	Asp	Leu	Met	Leu	Asn	Ile
			535					540						545	
Glu	Ser	Ala	Ile	Lys	Lys	His	Val	Lys	Glu	Val	Lys	Ser	Ile	Val	Ala
			550					555					560		
Arg	Thr	Gly	Ser	Asp	Glu	Leu	Gly	Leu	Asp	Leu	Gly	Gly	Leu	Asn	Gln
		565					570					575			
Thr	Asp	Thr	Phe	Ile	Ser	Phe	Ile	Pro	Lys	Lys	Glu	Trp	Ser	Val	Lys
	580				585										

-233-

```

Gly Val Asn Tyr Leu Tyr Val Thr Pro Asn Lys Glu Ser Met Ala Asp
675                      680                      685                      690
Val Gly Ile Thr Ser Asp Glu Phe Ser Lys Phe Leu Lys Ser Ala Leu
                      695                      700                      705
Glu Gly Leu Val Val Asp Val Ile Pro Thr Gly Ile Ser Arg Thr Pro
                      710                      715                      720
Val Met Ile Arg Gln Glu Ser Asp Phe Ala Ser Ser Ile Thr Lys Ile
                      725                      730                      735
Lys Ser Leu Ala Leu Thr Ser Lys Tyr Gly Val Leu Val Pro Ile Thr
                      740                      745                      750
Ser Ile Ala Lys Ile Glu Val Asp Gly Pro Val Ser Val Val Arg
755                      760                      765                      770
Glu Asn Ser Met Arg Met Ser Val Val Arg Ser Asn Val Val Gly Arg
                      775                      780                      785
Asp Leu Lys Ser Phe Val Glu Glu Ala Lys Lys Val Ile Ala Gln Asn
                      790                      795                      800
Ile Lys Leu Pro Pro Ser Tyr Tyr Ile Thr Tyr Gly Gly Gln Phe Glu
                      805                      810                      815
Asn Gln Gln Arg Ala Asn Lys Arg Leu Ser Thr Val Ile Pro Leu Ser
                      820                      825                      830
Ile Leu Ala Ile Phe Phe Ile Leu Phe Phe Thr Phe Lys Ser Ile Pro
835                      840                      845                      850
Leu Ala Leu Leu Ile Leu Leu Asn Ile Pro Phe Ala Val Thr Gly Gly
                      855                      860                      865
Leu Ile Ala Leu Phe Ala Val Gly Glu Tyr Ile Ser Val Pro Ala Ser
                      870                      875                      880
Val Gly Phe Ile Ala Leu Phe Gly Ile Ala Val Leu Asn Gly Val Val
                      885                      890                      895
Met Ile Gly Tyr Phe Lys Glu Leu Leu Leu Gln Gly Lys Ser Val Glu
900                      905                      910
Glu Cys Val Leu Leu Gly Ala Lys Arg Arg Leu Arg Pro Val Leu Met
915                      920                      925                      930
Thr Ala Cys Ile Ala Gly Leu Gly Leu Leu Pro Leu Leu Phe Ser His
                      935                      940                      945
Ser Val Gly Ser Glu Val Gln Lys Pro Leu Ala Ile Val Val Leu Gly
                      950                      955                      960
Gly Leu Val Thr Ser Ser Ala Leu Thr Leu Leu Leu Leu Pro Pro Met
965                      970                      975
Phe Met Leu Ile Ala Lys Lys Ile Lys Ile Val
980                      985

```

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 898 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 86...835
- (D) OTHER INFORMATION:

-234-

- (A) NAME/KEY: Signal Sequence
 (B) LOCATION: 86...161
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

GCATAAAATA AACAAACATT AAGTAAGGCT TATCAATATT TGATTACAAT TATAAGGGTT	60
ACATTTTTTT AATAGGAGAT ATACC ATG CTA GGA AAC GTT AAA AAA ACC CTT	112
Met Leu Gly Asn Val Lys Lys Thr Leu	
-25 -20	
TTT GGG GTC TTG TGT TTG GGC ACG TTG TGT TTG AGA GGG TTA ATG GCA	160
Phe Gly Val Leu Cys Leu Gly Thr Leu Cys Leu Arg Gly Leu Met Ala	
-15 -10 -5	
GAG CCA GAC GCT AAA GAG CTT GTT AAT TTA GGC ATA GAG AGC GCG AAG	208
Glu Pro Asp Ala Lys Glu Leu Val Asn Leu Gly Ile Glu Ser Ala Lys	
1 5 10 15	
AAG CAA GAT TTC GCT CAA GCT AAA ACG CAT TTT GAA AAA GCT TGT GAG	256
Lys Gln Asp Phe Ala Gln Ala Lys Thr His Phe Glu Lys Ala Cys Glu	
20 25 30	
TTA AAA AAT GGC TTT GGA TGT GTT TTT TTA GGG GCG TTC TAT GAA GAA	304
Leu Lys Asn Gly Phe Gly Cys Val Phe Leu Gly Ala Phe Tyr Glu Glu	
35 40 45	
GGG AAA GGA GTG GGA AAA GAC TTG AAA AAA GCC ATC CAA TTT TAC ACT	352
Gly Lys Gly Val Gly Lys Asp Leu Lys Lys Ala Ile Gln Phe Tyr Thr	
50 55 60	
AAA GGT TGT GAA TTA AAT GAT GGT TAT GGG TGT AAC CTG CTA GGA AAT	400
Lys Gly Cys Glu Leu Asn Asp Gly Tyr Gly Cys Asn Leu Leu Gly Asn	
65 70 75 80	
TTA TAC TAT AAC GGA CAA GGC GTG TCA AAA GAC GCT AAA AAA GCC TCA	448
Leu Tyr Tyr Asn Gly Gln Gly Val Ser Lys Asp Ala Lys Lys Ala Ser	
85 90 95	
CAA TAC TAC TCT AAA GCT TGC GAC TTA AAC CAT GCT GAA GGG TGT ATG	496
Gln Tyr Tyr Ser Lys Ala Cys Asp Leu Asn His Ala Glu Gly Cys Met	
100 105 110	
GTA TTA GGA AGC TTA CAC CAT TAT GGC GTA GGC ACG CCT AAG GAT TTA	544
Val Leu Gly Ser Leu His His Tyr Gly Val Gly Thr Pro Lys Asp Leu	
115 120 125	
AGA AAG GCT CTT GAT TTG TAT GAA AAA GCT TGC GAT TTA AAA GAC AGC	592
Arg Lys Ala Leu Asp Leu Tyr Glu Lys Ala Cys Asp Leu Lys Asp Ser	
130 135 140	

CCA	GGG	TGT	ATT	AAT	GCA	GGA	TAT	ATA	TAT	AGT	GTA	ACA	AAG	AAT	TTT	640
Pro	Gly	Cys	Ile	Asn	Ala	Gly	Tyr	Ile	Tyr	Ser	Val	Thr	Lys	Asn	Phe	
145					150					155					160	
AAG	GAG	GCT	ATC	GTT	CGT	TAT	TCT	AAA	GCA	TGC	GAA	TTA	AAA	GAT	GGT	688
Lys	Glu	Ala	Ile	Val	Arg	Tyr	Ser	Lys	Ala	Cys	Glu	Leu	Lys	Asp	Gly	
				165					170					175		
AGG	GGG	TGT	TAT	AAT	TTA	GGG	GTT	ATG	CAA	TAC	AAC	GCT	CAA	GGT	ACA	736
Arg	Gly	Cys	Tyr	Asn	Leu	Gly	Val	Met	Gln	Tyr	Asn	Ala	Gln	Gly	Thr	
			180					185					190			
GCA	AAG	GAC	GAA	AAG	CAA	GCG	GTA	GAA	AAC	TTT	AAA	AAA	GGC	TGC	AAA	784
Ala	Lys	Asp	Glu	Lys	Gln	Ala	Val	Glu	Asn	Phe	Lys	Lys	Gly	Cys	Lys	
		195					200					205				
TCA	AGC	GTT	AAA	GAA	GCA	TGC	GAC	GCT	CTC	AAG	GAA	TTA	AAA	ATA	GAA	832
Ser	Ser	Val	Lys	Glu	Ala	Cys	Asp	Ala	Leu	Lys	Glu	Leu	Lys	Ile	Glu	
	210					215					220					
CTT	TAATTTCAAT	GAAGTTAGCT	AAACGCTGCG	TTTAGCTGGC	TTTTACGCTT	TTTATA	891									
Leu																
225																
TTTTAAG																898

(A) LENGTH: 250 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
(B) LOCATION: 1...25
(D) OTHER INFORMATION:

Met	Leu	Gly	Asn	Val	Lys	Lys	Thr	Leu	Phe	Gly	Val	Leu	Cys	Leu	Gly
-25					-20					-15					-10
Thr	Leu	Cys	Leu	Arg	Gly	Leu	Met	Ala	Glu	Pro	Asp	Ala	Lys	Glu	Leu
				-5					1				5		
Val	Asn	Leu	Gly	Ile	Glu	Ser	Ala	Lys	Lys	Gln	Asp	Phe	Ala	Gln	Ala
	10						15				20				
Lys	Thr	His	Phe	Glu	Lys	Ala	Cys	Glu	Leu	Lys	Asn	Gly	Phe	Gly	Cys
	25					30					35				
Val	Phe	Leu	Gly	Ala	Phe	Tyr	Glu	Glu	Gly	Lys	Gly	Val	Gly	Lys	Asp
40					45					50					55

-236-

```

Leu Lys Lys Ala Ile Gln Phe Tyr Thr Lys Gly Cys Glu Leu Asn Asp
      60                      65                      70
Gly Tyr Gly Cys Asn Leu Leu Gly Asn Leu Tyr Tyr Asn Gly Gln Gly
      75                      80                      85
Val Ser Lys Asp Ala Lys Lys Ala Ser Gln Tyr Tyr Ser Lys Ala Cys
      90                      95                      100
Asp Leu Asn His Ala Glu Gly Cys Met Val Leu Gly Ser Leu His His
      105                     110                     115
Tyr Gly Val Gly Thr Pro Lys Asp Leu Arg Lys Ala Leu Asp Leu Tyr
      120                     125                     130                     135
Glu Lys Ala Cys Asp Leu Lys Asp Ser Pro Gly Cys Ile Asn Ala Gly
      140                     145                     150
Tyr Ile Tyr Ser Val Thr Lys Asn Phe Lys Glu Ala Ile Val Arg Tyr
      155                     160                     165
Ser Lys Ala Cys Glu Leu Lys Asp Gly Arg Gly Cys Tyr Asn Leu Gly
      170                     175                     180
Val Met Gln Tyr Asn Ala Gln Gly Thr Ala Lys Asp Glu Lys Gln Ala
      185                     190                     195
Val Glu Asn Phe Lys Lys Gly Cys Lys Ser Ser Val Lys Glu Ala Cys
      200                     205                     210                     215
Asp Ala Leu Lys Glu Leu Lys Ile Glu Leu
      220                     225

```

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1079 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 169...834
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 169...289
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

```

CAAAAAAAAA AAAAAACAAT TTCAGTTTCT TATTAGCTAG GTTTGATTAA AATGAAAAGC      60
TTTTATGTGT TTAAACTTCA TTGTCTTAAA ACTTTTAAGA GCAATTTTAA AATTCGTTGG      120
CGTATAATAT CCGTTTTGAA TGAAC TACTA AAAAAAGGGT TTAAATA ATG GCT GAA      177
                                   Met Ala Glu
                                   -40

```

```

AAT TCT TTC AAA AAT GTT TCC ACA CAA CCC AAA GTA TTT TTC TTA TTG      225

```

-237-

Asn	Ser	Phe	Lys	Asn	Val	Ser	Thr	Gln	Pro	Lys	Val	Phe	Phe	Leu	Leu		
		-35					-30					-25					
CCA	GCT	AAA	ACC	CTG	TTT	CTT	TTA	GGA	GGC	GTT	TTT	AGC	GCG	TTT	TTT	273	
Pro	Ala	Lys	Thr	Leu	Phe	Leu	Leu	Gly	Gly	Val	Phe	Ser	Ala	Phe	Phe		
	-20					-15					-10						
ATC	CTT	ATT	GCT	GGC	TTG	GTT	TTT	TTT	GAT	TAT	GCT	CAT	TTG	ATG	GAC	321	
Ile	Leu	Ile	Ala	Gly	Leu	Val	Phe	Phe	Asp	Tyr	Ala	His	Leu	Met	Asp		
-5					1				5					10			
AAT	GCC	ATT	TTT	AAT	TTT	GCG	CGT	TCA	ACC	CCC	TTT	AAT	TCC	AGC	CCT	369	
Asn	Ala	Ile	Phe	Asn	Phe	Ala	Arg	Ser	Thr	Pro	Phe	Asn	Ser	Ser	Pro		
		15					20					25					
ATT	TTA	ACT	CTA	ATC	CTC	CAA	AAT	ATC	GCT	AAT	TTA	GGC	TCT	TCT	CAA	417	
Ile	Leu	Thr	Leu	Ile	Leu	Gln	Asn	Ile	Ala	Asn	Leu	Gly	Ser	Ser	Gln		
		30					35					40					
TTC	GTG	TTG	CCT	TTG	AGT	TTG	TTG	GTG	GGG	GTG	TTT	TTA	AGC	CTT	TAT	465	
Phe	Val	Leu	Pro	Leu	Ser	Leu	Leu	Val	Gly	Val	Phe	Leu	Ser	Leu	Tyr		
	45					50					55						
CGC	AGA	AAC	TTA	GTG	CTT	GGG	GTG	TGG	TTT	GTG	TTA	AGC	GTG	ATC	TTG	513	
Arg	Arg	Asn	Leu	Val	Leu	Gly	Val	Trp	Phe	Val	Leu	Ser	Val	Ile	Leu		
60					65					70					75		
TTT	GAA	GCC	CTT	TTA	GAA	TCT	TTA	AAA	CAC	CTT	TTT	GCA	TAT	TCC	ATT	561	
Phe	Glu	Ala	Leu	Leu	Glu	Ser	Leu	Lys	His	Leu	Phe	Ala	Tyr	Ser	Ile		
				80					85					90			
CAG	TGG	CTT	TCG	CGC	AGC	GCT	AAT	TTC	CCT	AAC	GCT	ACT	GCG	CTT	TCT	609	
Gln	Trp	Leu	Ser	Arg	Ser	Ala	Asn	Phe	Pro	Asn	Ala	Thr	Ala	Leu	Ser		
			95					100					105				
TTA	GTG	CTA	TTT	TAT	GGG	TTG	CTT	ATT	TTA	TTG	ATA	CCC	CAT	TTA	ATC	657	
Leu	Val	Leu	Phe	Tyr	Gly	Leu	Leu	Ile	Leu	Leu	Ile	Pro	His	Leu	Ile		
		110					115					120					
ACG	CAT	CAA	ACG	CTT	AAA	AAT	GTT	CTT	TTT	TAT	AGC	TTA	TTT	GGT	TTG	705	
Thr	His	Gln	Thr	Leu	Lys	Asn	Val	Leu	Phe	Tyr	Ser	Leu	Phe	Gly	Leu		
	125					130					135						
ATT	TTT	TTA	ATA	GGG	TTA	GCA	CTG	ATT	GTT	TTA	GGG	GTT	TCT	TTC	AGT	753	
Ile	Phe	Leu	Ile	Gly	Leu	Ala	Leu	Ile	Val	Leu	Gly	Val	Ser	Phe	Ser		
140					145					150					155		
AGT	GTT	TTA	GGA	GGG	TTT	TGT	TTA	GGG	GCG	TTA	GGG	GCT	TGT	TTT	TCC	801	
Ser	Val	Leu	Gly	Gly	Phe	Cys	Leu	Gly	Ala	Leu	Gly	Ala	Cys	Phe	Ser		
				160					165					170			
ATA	GGG	ATT	TAT	TTG	AGC	GTG	TTT	CAA	AAG	ATC	TAAACGAACG	GCTTAAAAGA				854	
Ile	Gly	Ile	Tyr	Leu	Ser	Val	Phe	Gln	Lys	Ile							
			175					180									
ATGAAAATTT	TATCAAGGTT	TTAATATTGG	ATTTAAAGGT	ATTATTGCAA	CGGATTGTTG											914	

-238-

```

ATTTTTCAT CAAGCTCAAT AAAAAGCAAA AAATCGCCCT GATTGCAGCT GGGGTTTTGA      974
TCACGGCTTT GCTTGTGTTT TTATTGCTCT ATCCCTTTAA AGAAAAAGAC TACACGCAAG      1034
GGGGTTATGG GGTTTTATTT GAAGGTTTAG ACTCTAGCGA TAACG                      1079

```

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

```

Met Ala Glu Asn Ser Phe Lys Asn Val Ser Thr Gln Pro Lys Val Phe
-40          -35          -30          -25
Phe Leu Leu Pro Ala Lys Thr Leu Phe Leu Leu Gly Gly Val Phe Ser
          -20          -15          -10
Ala Phe Phe Ile Leu Ile Ala Gly Leu Val Phe Phe Asp Tyr Ala His
          -5           1           5
Leu Met Asp Asn Ala Ile Phe Asn Phe Ala Arg Ser Thr Pro Phe Asn
 10          15          20
Ser Ser Pro Ile Leu Thr Leu Ile Leu Gln Asn Ile Ala Asn Leu Gly
25          30          35          40
Ser Ser Gln Phe Val Leu Pro Leu Ser Leu Leu Val Gly Val Phe Leu
          45          50          55
Ser Leu Tyr Arg Arg Asn Leu Val Leu Gly Val Trp Phe Val Leu Ser
          60          65          70
Val Ile Leu Phe Glu Ala Leu Leu Glu Ser Leu Lys His Leu Phe Ala
          75          80          85
Tyr Ser Ile Gln Trp Leu Ser Arg Ser Ala Asn Phe Pro Asn Ala Thr
 90          95          100
Ala Leu Ser Leu Val Leu Phe Tyr Gly Leu Leu Ile Leu Leu Ile Pro
105          110          115          120
His Leu Ile Thr His Gln Thr Leu Lys Asn Val Leu Phe Tyr Ser Leu
          125          130          135
Phe Gly Leu Ile Phe Leu Ile Gly Leu Ala Leu Ile Val Leu Gly Val
          140          145          150
Ser Phe Ser Ser Val Leu Gly Gly Phe Cys Leu Gly Ala Leu Gly Ala
155          160          165
Cys Phe Ser Ile Gly Ile Tyr Leu Ser Val Phe Gln Lys Ile
170          175          180

```

(2) INFORMATION FOR SEQ ID NO:113:

-239-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 962 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 97...912
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 97...217
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

TTTTATTGAA	TGTGTTGTAA	TGTTTTTAAG	GTATAATAAA	CTCTTTTAA	GTCAAGCAAT	60
AAAGTTTGCA	ACCTGATGAG	AGTAATAATA	GAGTTT ATG	CTG ATT TCA	TTA AAA	114
			Met	Leu Ile Ser	Leu Lys	
			-40		-35	
ACA TTC CTA	AAA ATA TTA	TTG AAA ATA	TTC CTA AAA	ACC TTC CAA	AAG	162
Thr Phe Leu	Lys Ile Leu	Leu Lys Ile	Phe Leu Lys	Thr Phe Gln	Lys	
	-30		-25		-20	
ATT TGG GTA	GTT TGC GTT	ATT ATT TGG	GGG TTA GGC	TGT AGT TTT	TTA	210
Ile Trp Val	Val Cys Val	Ile Ile Trp	Gly Leu Gly	Cys Ser Phe	Leu	
	-15		-10		-5	
AAC GCT AAC	AGC ATT CAA	TTA GAA GAA	ACG CTC AGA	CGA AGC CCT	AAA	258
Asn Ala Asn	Ser Ile Gln	Leu Glu Glu	Thr Leu Arg	Arg Ser Pro	Lys	
	1	5	10			
AAT CTT ATT	TGG CAA CAC	TTT AAA AAG	AAG TTT AAA	AAG AGC AAC	ACG	306
Asn Leu Ile	Trp Gln His	Phe Lys Lys	Lys Phe Lys	Lys Ser Asn	Thr	
15	20	25	30			
ATC CCT TAT	GCC CCA AAT	AGC CGT TGG	AAA TAT TTA	GGC ACG AGC	ATA	354
Ile Pro Tyr	Ala Pro Asn	Ser Arg Trp	Lys Tyr Leu	Gly Thr Ser	Ile	
	35	40	45			
GGG ATT TTA	GGC GTG TCT	TTG GTG ATA	GGG ATT GTG	GGG CTG TAT	CTC	402
Gly Ile Leu	Gly Val Ser	Leu Val Ile	Gly Ile Val	Gly Leu Tyr	Leu	
	50	55	60			
ATG CCA GAG	AGC GTA ACG	AAT TGG GAT	AAA GAA AAG	TTT GGG ATC	AAA	450
Met Pro Glu	Ser Val Thr	Asn Trp Asp	Lys Glu Lys	Phe Gly Ile	Lys	
	65	70	75			

-240-

AGT TGG TTT GAA AAT GTC CGC ATG GGG CCA AAA CTG GAC AAT GAT AGT	498
Ser Trp Phe Glu Asn Val Arg Met Gly Pro Lys Leu Asp Asn Asp Ser	
80 85 90	
TTT ATT TTT AAT GAA ATT TTG CAC CCT TAT TTT GGG GCT ATG TAT TAT	546
Phe Ile Phe Asn Glu Ile Leu His Pro Tyr Phe Gly Ala Met Tyr Tyr	
95 100 105 110	
ATG CAA CCG CGC ATG GCT GGA TTT AGC TGG ATG GCA TCA GCG TTT TTT	594
Met Gln Pro Arg Met Ala Gly Phe Ser Trp Met Ala Ser Ala Phe Phe	
115 120 125	
TCT TTT ATC ACT TCC ACG CTT TTT TGG GAA TAT GGC TTG GAA GCG TTT	642
Ser Phe Ile Thr Ser Thr Leu Phe Trp Glu Tyr Gly Leu Glu Ala Phe	
130 135 140	
GTG GAA GTG CCT AGC TGG CAG GAT TTA GTG ATC ACG CCT TTA TTA GGC	690
Val Glu Val Pro Ser Trp Gln Asp Leu Val Ile Thr Pro Leu Leu Gly	
145 150 155	
TCC ATT TTA GGG GAG GGG TTT TAT CAG CTC ACG CGC TAT ATC CAA CGC	738
Ser Ile Leu Gly Glu Gly Phe Tyr Gln Leu Thr Arg Tyr Ile Gln Arg	
160 165 170	
AAT GAA GGC AAG CTT TTT GGC TCT TTA TTT TTA GGG CGT TTA GTC ATC	786
Asn Glu Gly Lys Leu Phe Gly Ser Leu Phe Leu Gly Arg Leu Val Ile	
175 180 185 190	
GCT CTT ATG GAT CCT ATC GGT TTT ATC ATT AGG GAT TTA GGA CTT GGG	834
Ala Leu Met Asp Pro Ile Gly Phe Ile Ile Arg Asp Leu Gly Leu Gly	
195 200 205	
GAA GCT TTA GGG ATT TAT AAT AAA CAC GAA ATC CGT TCC AGC TTA AGC	882
Glu Ala Leu Gly Ile Tyr Asn Lys His Glu Ile Arg Ser Ser Leu Ser	
210 215 220	
CCC AAT GGT TTG AAT TTG ACT TAC AAA TTT TAAGAGCTTA AAATTTAAGA AAA	935
Pro Asn Gly Leu Asn Leu Thr Tyr Lys Phe	
225 230	
TTATAAAGAG TTTTGATAGA ATACCTT	962

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence

-241-

(B) LOCATION: 1...40

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

```

Met Leu Ile Ser Leu Lys Thr Phe Leu Lys Ile Leu Leu Lys Ile Phe
-40          -35          -30          -25
Leu Lys Thr Phe Gln Lys Ile Trp Val Val Cys Val Ile Ile Trp Gly
          -20          -15          -10
Leu Gly Cys Ser Phe Leu Asn Ala Asn Ser Ile Gln Leu Glu Glu Thr
          -5          1          5
Leu Arg Arg Ser Pro Lys Asn Leu Ile Trp Gln His Phe Lys Lys Lys
10          15          20
Phe Lys Lys Ser Asn Thr Ile Pro Tyr Ala Pro Asn Ser Arg Trp Lys
25          30          35          40
Tyr Leu Gly Thr Ser Ile Gly Ile Leu Gly Val Ser Leu Val Ile Gly
          45          50          55
Ile Val Gly Leu Tyr Leu Met Pro Glu Ser Val Thr Asn Trp Asp Lys
60          65          70
Glu Lys Phe Gly Ile Lys Ser Trp Phe Glu Asn Val Arg Met Gly Pro
75          80          85
Lys Leu Asp Asn Asp Ser Phe Ile Phe Asn Glu Ile Leu His Pro Tyr
90          95          100
Phe Gly Ala Met Tyr Tyr Met Gln Pro Arg Met Ala Gly Phe Ser Trp
105          110          115          120
Met Ala Ser Ala Phe Phe Ser Phe Ile Thr Ser Thr Leu Phe Trp Glu
125          130          135
Tyr Gly Leu Glu Ala Phe Val Glu Val Pro Ser Trp Gln Asp Leu Val
140          145          150
Ile Thr Pro Leu Leu Gly Ser Ile Leu Gly Glu Gly Phe Tyr Gln Leu
155          160          165
Thr Arg Tyr Ile Gln Arg Asn Glu Gly Lys Leu Phe Gly Ser Leu Phe
170          175          180
Leu Gly Arg Leu Val Ile Ala Leu Met Asp Pro Ile Gly Phe Ile Ile
185          190          195          200
Arg Asp Leu Gly Leu Gly Glu Ala Leu Gly Ile Tyr Asn Lys His Glu
205          210          215
Ile Arg Ser Ser Leu Ser Pro Asn Gly Leu Asn Leu Thr Tyr Lys Phe
220          225          230

```

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1422 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 216...1202

-242-

(D) OTHER INFORMATION:

(A) NAME/KEY: Signal Sequence

(B) LOCATION: 216...273

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

```

AAATTAAACG AGTTTGTTT AGAGCCGTAT TTAGGGTTTT TGCACCCCCA TTAAACCAAT      60
GATTTTGAAA ATAACCCTAA TGAGCAATCA GCGCTCTTTG TCTTGCCCCT TTCAGCGGTT      120
AGCGCTCTTA ATGTGCATGC ACTCAAATTT GTGTTGTTGG AAGCGTTACC CTAAAACGCT      180
ATTTTAAAAA TAATCCATTA AAATAAAGGC GAGGA ATG AAA AGA TTT GTT TTG      233
                               Met Lys Arg Phe Val Leu
                               -15

TTT TTA TTG TTC ATG TGC GTT TGC GTT CAA GCT TAC GCC GAG CAA GAT      281
Phe Leu Leu Phe Met Cys Val Cys Val Gln Ala Tyr Ala Glu Gln Asp
                               -10                               -5                               1

TAC TTT TTT AGG GAT TTT AAA TCT AGA GAT TTG CCC CAA AAA CTC CAT      329
Tyr Phe Phe Arg Asp Phe Lys Ser Arg Asp Leu Pro Gln Lys Leu His
                               5                               10                               15

CTT GAT AAA AAG CTC TCC CAA ACA ATA CAG CCA TGC ATG CAA CTT AAC      377
Leu Asp Lys Lys Leu Ser Gln Thr Ile Gln Pro Cys Met Gln Leu Asn
20                               25                               30                               35

GCA TCA AAA CAC TAC ACT TCT ACC GGG GTT AGA GAG CCT GAT AAA TGC      425
Ala Ser Lys His Tyr Thr Ser Thr Gly Val Arg Glu Pro Asp Lys Cys
                               40                               45                               50

ACA AAG AGT TTT AAA AAA TCC GCT CTC ATG TCC TAT GAC TTA GCG CTA      473
Thr Lys Ser Phe Lys Lys Ser Ala Leu Met Ser Tyr Asp Leu Ala Leu
                               55                               60                               65

GGT TAT TTG GTG AGT AAG AAT AAG CAA TAC GGC TTA AAG GCT ATA GAA      521
Gly Tyr Leu Val Ser Lys Asn Lys Gln Tyr Gly Leu Lys Ala Ile Glu
70                               75                               80

ATT TTA AAC GCT TGG GCT AAA GAG CTT CAA AGC GTG GAT ACT TAT CAG      569
Ile Leu Asn Ala Trp Ala Lys Glu Leu Gln Ser Val Asp Thr Tyr Gln
85                               90                               95

AGC GAG GAT AAT ATC AAT TTT TAC ATG CCT TAT ATG AAC ATG GCT TAT      617
Ser Glu Asp Asn Ile Asn Phe Tyr Met Pro Tyr Met Asn Met Ala Tyr
100                               105                               110                               115

TGG TTT GTC AAA AAG GCG TTT CCT AGC CCA GAA TAT GAA GAT TTC ATT      665
Trp Phe Val Lys Lys Ala Phe Pro Ser Pro Glu Tyr Glu Asp Phe Ile
                               120                               125                               130

AAG CGG ATG CGC CAG TAT TCT CAA TCA GCT CTT AAC ACT AAC CAT GGG      713

```

-243-

Lys	Arg	Met	Arg	Gln	Tyr	Ser	Gln	Ser	Ala	Leu	Asn	Thr	Asn	His	Gly		
			135					140					145				
GCG	TGG	GGC	ATT	CTT	TTT	GAT	GTG	AGT	TCT	GCG	CTA	GCG	TTA	GAC	GAT	761	
Ala	Trp	Gly	Ile	Leu	Phe	Asp	Val	Ser	Ser	Ala	Leu	Ala	Leu	Asp	Asp		
		150					155					160					
AAT	GCC	CTT	TTG	CAC	AAT	AGC	GCT	AAT	CGG	TGG	CAG	GAG	TGG	GTG	TTT	809	
Asn	Ala	Leu	Leu	His	Asn	Ser	Ala	Asn	Arg	Trp	Gln	Glu	Trp	Val	Phe		
	165					170				175							
AAA	GCC	ATA	GAT	GAG	AAT	GGG	GTT	ATT	GNT	AGC	GCG	ATC	ACT	AGG	AGC	857	
Lys	Ala	Ile	Asp	Glu	Asn	Gly	Val	Ile	Xaa	Ser	Ala	Ile	Thr	Arg	Ser		
180					185					190					195		
GAT	ACG	AGC	GAT	TAT	CAT	GGC	GGC	CCT	ACA	AAG	GGC	ATT	AAG	GGG	ATA	905	
Asp	Thr	Ser	Asp	Tyr	His	Gly	Gly	Pro	Thr	Lys	Gly	Ile	Lys	Gly	Ile		
				200					205					210			
GCT	TAT	ACC	AAT	TTC	GCG	CTT	CTT	GCG	CTA	ACC	ATA	TCA	GCG	GAA	TTG	953	
Ala	Tyr	Thr	Asn	Phe	Ala	Leu	Leu	Ala	Leu	Thr	Ile	Ser	Gly	Glu	Leu		
			215					220					225				
CTT	TTT	GAG	AAC	GGG	TAT	GAT	TTG	TGG	GGT	AGT	GGA	GCT	GGG	AAA	AGG	1001	
Leu	Phe	Glu	Asn	Gly	Tyr	Asp	Leu	Trp	Gly	Ser	Gly	Ala	Gly	Lys	Arg		
		230					235					240					
CTC	TCT	GTG	GCG	TAT	AAC	AAA	GTT	GCA	ACA	TGG	ATT	TTA	AAC	CCT	GAA	1049	
Leu	Ser	Val	Ala	Tyr	Asn	Lys	Val	Ala	Thr	Trp	Ile	Leu	Asn	Pro	Glu		
	245					250					255						
ACT	TTC	CCT	TAT	TTC	CAG	CCT	AAC	CTT	ATC	GGG	GTG	CAT	AAC	AAC	GCC	1097	
Thr	Phe	Pro	Tyr	Phe	Gln	Pro	Asn	Leu	Ile	Gly	Val	His	Asn	Asn	Ala		
260					265				270						275		
TAT	TTC	ATT	ATT	TTA	GCC	AAG	CAT	TAT	TCT	AGC	CCT	AGT	GCA	AAT	GAG	1145	
Tyr	Phe	Ile	Ile	Leu	Ala	Lys	His	Tyr	Ser	Ser	Pro	Ser	Ala	Asn	Glu		
				280					285					290			
CTT	TTA	AAG	CAA	GGC	GAT	TTA	CAC	GAA	GAT	GGT	TTC	AGG	CTG	AAA	CTC	1193	
Leu	Leu	Lys	Gln	Gly	Asp	Leu	His	Glu	Asp	Gly	Phe	Arg	Leu	Lys	Leu		
		295						300				305					
CGA	TCG	CCA	TGAATTTTTC	TGTATCCAAG	GTTAGCCTTA	AGGATGGCCA	TGCGCTTTA	1251									
Arg	Ser	Pro															
		310															
ACCTTTTGAT	GAATGGTTCA	GAAAGTTTGT	TTCAGTCAGC	ATTATTTTACA	AAAAGAGTTT	1311											
AAAATAAACG	CAATTGTATC	TCTTGAGTCG	TCTTTAGAGT	GCAAAATGATT	ATCAAAATGA	1371											
ATCGTTTTAG	TTGTAAGCGT	GCTTATTTAC	ACTAAAATAA	TAAGCGTTAT	T	1422											

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 329 amino acids

-244-

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal
 (ix) FEATURE:

(A) NAME/KEY: Signal Sequence
 (B) LOCATION: 1...19
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

```

Met Lys Arg Phe Val Leu Phe Leu Leu Phe Met Cys Val Cys Val Gln
              -15                      -10                      -5
Ala Tyr Ala Glu Gln Asp Tyr Phe Phe Arg Asp Phe Lys Ser Arg Asp
              1                      5                      10
Leu Pro Gln Lys Leu His Leu Asp Lys Lys Leu Ser Gln Thr Ile Gln
15                      20                      25
Pro Cys Met Gln Leu Asn Ala Ser Lys His Tyr Thr Ser Thr Gly Val
30                      35                      40                      45
Arg Glu Pro Asp Lys Cys Thr Lys Ser Phe Lys Lys Ser Ala Leu Met
              50                      55                      60
Ser Tyr Asp Leu Ala Leu Gly Tyr Leu Val Ser Lys Asn Lys Gln Tyr
              65                      70                      75
Gly Leu Lys Ala Ile Glu Ile Leu Asn Ala Trp Ala Lys Glu Leu Gln
              80                      85                      90
Ser Val Asp Thr Tyr Gln Ser Glu Asp Asn Ile Asn Phe Tyr Met Pro
95                      100                      105
Tyr Met Asn Met Ala Tyr Trp Phe Val Lys Lys Ala Phe Pro Ser Pro
110                      115                      120                      125
Glu Tyr Glu Asp Phe Ile Lys Arg Met Arg Gln Tyr Ser Gln Ser Ala
              130                      135                      140
Leu Asn Thr Asn His Gly Ala Trp Gly Ile Leu Phe Asp Val Ser Ser
              145                      150                      155
Ala Leu Ala Leu Asp Asp Asn Ala Leu Leu His Asn Ser Ala Asn Arg
160                      165                      170
Trp Gln Glu Trp Val Phe Lys Ala Ile Asp Glu Asn Gly Val Ile Xaa
175                      180                      185
Ser Ala Ile Thr Arg Ser Asp Thr Ser Asp Tyr His Gly Gly Pro Thr
190                      195                      200                      205
Lys Gly Ile Lys Gly Ile Ala Tyr Thr Asn Phe Ala Leu Leu Ala Leu
              210                      215                      220
Thr Ile Ser Gly Glu Leu Leu Phe Glu Asn Gly Tyr Asp Leu Trp Gly
              225                      230                      235
Ser Gly Ala Gly Lys Arg Leu Ser Val Ala Tyr Asn Lys Val Ala Thr
240                      245                      250
Trp Ile Leu Asn Pro Glu Thr Phe Pro Tyr Phe Gln Pro Asn Leu Ile
255                      260                      265
Gly Val His Asn Asn Ala Tyr Phe Ile Ile Leu Ala Lys His Tyr Ser
270                      275                      280                      285
Ser Pro Ser Ala Asn Glu Leu Leu Lys Gln Gly Asp Leu His Glu Asp

```

-245-

	290	295	300
Gly Phe Arg Leu Lys Leu Arg Ser Pro			
	305	310	

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1080 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 157...987
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 157...226
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

AGCGGTAAAA TCGCTGAAGA AAACAACGCT AAAGAATTTT TTAACCACCC GAAATCTCAA	60
AGAGCGCAAA AATTTT TAGA AACTTTCCAT TTTT TAGGGA GCTGT TAAAT AAAGTTTGCT	120
AAAAAGATGA TTCTAATTTC AAAAAAAGGT GTTTTT ATG AAA ACA AAC GGG CTT	174
Met Lys Thr Asn Gly Leu	
-20	
TTT AAA ATG TGG GGG CTG TTT TTA GTT TTA ATC GCT TTA GTC TTT AAT	222
Phe Lys Met Trp Gly Leu Phe Leu Val Leu Ile Ala Leu Val Phe Asn	
-15 -10 -5	
GCA TGT TCT GAT AGC CAT AAA GAA AAA AAG GAC GCT TTA GAA GTC ATT	270
Ala Cys Ser Asp Ser His Lys Glu Lys Lys Asp Ala Leu Glu Val Ile	
1 5 10 15	
AAA CAA AGA GGG GTT TTA AAA GTG GGG GTT TTT AGC GAT AAG CCT CCT	318
Lys Gln Arg Gly Val Leu Lys Val Gly Val Phe Ser Asp Lys Pro Pro	
20 25 30	
TTT GGC TCT GTG GAT TCT AAA GGG AAA TAT CAA GGC TAT GAT GTA GTT	366
Phe Gly Ser Val Asp Ser Lys Gly Lys Tyr Gln Gly Tyr Asp Val Val	
35 40 45	
ATT GCT AAA CGC ATG GCT CTT GAT TTA TTG GGC GAT GAA AAT AAG ATT	414
Ile Ala Lys Arg Met Ala Leu Asp Leu Leu Gly Asp Glu Asn Lys Ile	
50 55 60	

-246-

GAG TTT ATT CCT GTA GAA GCT TCA GCT AGG GTG GAA TTT TTA AAA GCC Glu Phe Ile Pro Val Glu Ala Ser Ala Arg Val Glu Phe Leu Lys Ala 65 70 75	462
AAT AAA GTG GAT ATT ATC ATG GCT AAT TTC ACG CGC ACT AAA GAA AGA Asn Lys Val Asp Ile Ile Met Ala Asn Phe Thr Arg Thr Lys Glu Arg 80 85 90 95	510
GAA AAA GTC GTG GAT TTC GCT AAG CCG TAT ATG AAA GTC GCT TTA GGG Glu Lys Val Val Asp Phe Ala Lys Pro Tyr Met Lys Val Ala Leu Gly 100 105 110	558
GTG GTT TCT AAA GAT GGG GTC ATT AAA AAT ATA GAA GAG TTG AAA GAT Val Val Ser Lys Asp Gly Val Ile Lys Asn Ile Glu Glu Leu Lys Asp 115 120 125	606
AAA GAG TTG ATT GTG AAT AAA GGC ACG ACA GCG GAT TTT TAT TTC ACT Lys Glu Leu Ile Val Asn Lys Gly Thr Thr Ala Asp Phe Tyr Phe Thr 130 135 140	654
AAA AAT TAC CCC AAT ATC AAG CTT TTG AAA TTT GAG CAA AAT ACA GAG Lys Asn Tyr Pro Asn Ile Lys Leu Leu Lys Phe Glu Gln Asn Thr Glu 145 150 155	702
ACT TTT TTA GCC CTT TTA AAC AAT AAG GCT ACC GCT CTA GCC CAT GAC Thr Phe Leu Ala Leu Leu Asn Asn Lys Ala Thr Ala Leu Ala His Asp 160 165 170 175	750
AAC ACT TTA TTG CTC GCT TGG ACG AAA CAA CAC CCT GAA TTT AAA TTA Asn Thr Leu Leu Leu Ala Trp Thr Lys Gln His Pro Glu Phe Lys Leu 180 185 190	798
GGC ATT ACA AGC CTT GGC GAT AAG GAT GTG ATC GCT CCA GCG ATT AAA Gly Ile Thr Ser Leu Gly Asp Lys Asp Val Ile Ala Pro Ala Ile Lys 195 200 205	846
AAA GGC AAC CCC AAG CTT TTA GAA TGG TTG AAT AAC GAA ATA GAT TCC Lys Gly Asn Pro Lys Leu Leu Glu Trp Leu Asn Asn Glu Ile Asp Ser 210 215 220	894
CTC ATT TCT AGC GAC TTC TTA AAA GAA GCT TAT CAA GAG ACT TTA GCA Leu Ile Ser Ser Asp Phe Leu Lys Glu Ala Tyr Gln Glu Thr Leu Ala 225 230 235	942
CCT GTT TAT GGC GAT GAA ATC AAA CCG GAA GAA ATT ATT TTT GAA TGATT Pro Val Tyr Gly Asp Glu Ile Lys Pro Glu Glu Ile Ile Phe Glu 240 245 250	992
TCTTTAGGCT TTGAATTCTT GACAGGGTGC GTTTTTATTG CTAAATTAGC AATTTTGTGA TCTTTTGTGTT TTTCATTTTG AGATATAT	1052 1080

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 277 amino acids

-247-

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal
 (ix) FEATURE:

(A) NAME/KEY: Signal Sequence
 (B) LOCATION: 1...23
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

```

Met Lys Thr Asn Gly Leu Phe Lys Met Trp Gly Leu Phe Leu Val Leu
      -20                      -15                      -10
Ile Ala Leu Val Phe Asn Ala Cys Ser Asp Ser His Lys Glu Lys Lys
      -5                      1                      5
Asp Ala Leu Glu Val Ile Lys Gln Arg Gly Val Leu Lys Val Gly Val
10      15      20      25
Phe Ser Asp Lys Pro Pro Phe Gly Ser Val Asp Ser Lys Gly Lys Tyr
      30      35      40
Gln Gly Tyr Asp Val Val Ile Ala Lys Arg Met Ala Leu Asp Leu Leu
      45      50      55
Gly Asp Glu Asn Lys Ile Glu Phe Ile Pro Val Glu Ala Ser Ala Arg
      60      65      70
Val Glu Phe Leu Lys Ala Asn Lys Val Asp Ile Ile Met Ala Asn Phe
      75      80      85
Thr Arg Thr Lys Glu Arg Glu Lys Val Val Asp Phe Ala Lys Pro Tyr
90      95      100      105
Met Lys Val Ala Leu Gly Val Val Ser Lys Asp Gly Val Ile Lys Asn
      110      115      120
Ile Glu Glu Leu Lys Asp Lys Glu Leu Ile Val Asn Lys Gly Thr Thr
      125      130      135
Ala Asp Phe Tyr Phe Thr Lys Asn Tyr Pro Asn Ile Lys Leu Leu Lys
      140      145      150
Phe Glu Gln Asn Thr Glu Thr Phe Leu Ala Leu Leu Asn Asn Lys Ala
      155      160      165
Thr Ala Leu Ala His Asp Asn Thr Leu Leu Leu Ala Trp Thr Lys Gln
170      175      180      185
His Pro Glu Phe Lys Leu Gly Ile Thr Ser Leu Gly Asp Lys Asp Val
      190      195      200
Ile Ala Pro Ala Ile Lys Lys Gly Asn Pro Lys Leu Leu Glu Trp Leu
      205      210      215
Asn Asn Glu Ile Asp Ser Leu Ile Ser Ser Asp Phe Leu Lys Glu Ala
      220      225      230
Tyr Gln Glu Thr Leu Ala Pro Val Tyr Gly Asp Glu Ile Lys Pro Glu
      235      240      245
Glu Ile Ile Phe Glu
250

```

(2) INFORMATION FOR SEQ ID NO:119:

-248-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1114 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 37...1050
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

CGAGCTATCA CAACAAATCA ATTTGTAGGA ACAAGC ATG TTT TTT AAA ACT TAT	54
Met Phe Phe Lys Thr Tyr	
1 5	
CAA AAA TTA TTG GGT GCG AGC TGT TTG ACG TTG TAT TTA GCG GGC TGT	102
Gln Lys Leu Leu Gly Ala Ser Cys Leu Thr Leu Tyr Leu Ala Gly Cys	
10 15 20	
GGG AGT GAT AGT AGC GAG CCA TTG GTG GGA ATT GAA AAA AAT AGC TTC	150
Gly Ser Asp Ser Ser Glu Pro Leu Val Gly Ile Glu Lys Asn Ser Phe	
25 30 35	
AAT TCT ACC GTG AAA ATC ATT TCT AAA ACC GAC AAC ATA GAA ATC CAA	198
Asn Ser Thr Val Lys Ile Ile Ser Lys Thr Asp Asn Ile Glu Ile Gln	
40 45 50	
GAC TTG AAG CTC AAT CGT GGC AAT TGT GAG CAT GAT CAA AAT TTC TTG	246
Asp Leu Lys Leu Asn Arg Gly Asn Cys Glu His Asp Gln Asn Phe Leu	
55 60 65 70	
GTA AAG TTA ATC CAA GAA ACA GCC AAT ACA TAC CTG TTT GCA TCA GAA	294
Val Lys Leu Ile Gln Glu Thr Ala Asn Thr Tyr Leu Phe Ala Ser Glu	
75 80 85	
AAA GAA AAA GCG ATC AAA AAC CAC CAA GCA AAA ATC GCA AGA CTT CAA	342
Lys Glu Lys Ala Ile Lys Asn His Gln Ala Lys Ile Ala Arg Leu Gln	
90 95 100	
AAA GAT TTA GAA GAA CTC ACA CAG CAT GTG CAA CAA TCC AAT AAT CTT	390
Lys Asp Leu Glu Glu Leu Thr Gln His Val Gln Gln Ser Asn Asn Leu	
105 110 115	
GAT AAA TTG TTA GAA AAT GGA GGA CTA TTC GTT AGT GGC CAT GAT TAT	438
Asp Lys Leu Leu Glu Asn Gly Gly Leu Phe Val Ser Gly His Asp Tyr	
120 125 130	
AAA TAT ACA AAA GAT GAT AAC CCA ATA TAT GTT GTT AAG AGG ATG CTT	486
Lys Tyr Thr Lys Asp Asp Asn Pro Ile Tyr Val Val Lys Arg Met Leu	

-249-

135	140	145	150	
GAT AAC CTT GAT AGC TAT AAA TAT GAA TCA GAC GAC GTG CTA GAC GTG				534
Asp Asn Leu Asp Ser Tyr Lys Tyr Glu Ser Asp Asp Val Leu Asp Val	155	160	165	
CCA TAT GAG AAG CTA TTG GAA ATA AGC ATT GCT ATT GAA GAC ACT AAA				582
Pro Tyr Glu Lys Leu Leu Glu Ile Ser Ile Ala Ile Glu Asp Thr Lys	170	175	180	
AAC CCC AAA GAC TAC CCT TAT ATC AAC CTT AAA GAA CTC AAA AAA TTA				630
Asn Pro Lys Asp Tyr Pro Tyr Ile Asn Leu Lys Glu Leu Lys Lys Leu	185	190	195	
ATA GAT AGT ATT ATT GAT GAT CAT GGT TAT ATG GCC GAT GGC TTT TTG				678
Ile Asp Ser Ile Ile Asp Asp His Gly Tyr Met Ala Asp Gly Phe Leu	200	205	210	
AAT GAA TAT TCT AAT AGG GTA TCA AAA AAA GGT CTC CAA ATC CTT GCT				726
Asn Glu Tyr Ser Asn Arg Val Ser Lys Lys Gly Leu Gln Ile Leu Ala	215	220	225	230
AAA CTA AAA TCC ATG TGG CCT AGC GTA GGG AAA TTT TAT TTC GCC TCT				774
Lys Leu Lys Ser Met Trp Pro Ser Val Gly Lys Phe Tyr Phe Ala Ser	235	240	245	
TTG AAA GAG GCT ATC CCA AGG CAT GCC AAA GAA GTT ACT GAC AAG ATG				822
Leu Lys Glu Ala Ile Pro Arg His Ala Lys Glu Val Thr Asp Lys Met	250	255	260	
ATT AGC TCT GAA GAA AAA TCT ATC AAA GCC AAT CAA GTC AAA CTC ACT				870
Ile Ser Ser Glu Glu Lys Ser Ile Lys Ala Asn Gln Val Lys Leu Thr	265	270	275	
GAA GCG AAG CAA GAT ATT GAC AAA ATG GAA AAA ATC ATT AAA GAT TTA				918
Glu Ala Lys Gln Asp Ile Asp Lys Met Glu Lys Ile Ile Lys Asp Leu	280	285	290	
GAA AGC AAG AAA AAC ACC TTA TCA GTG TAT TTA AAA TTT GGA GAA AGT				966
Glu Ser Lys Lys Asn Thr Leu Ser Val Tyr Leu Lys Phe Gly Glu Ser	295	300	305	310
TTC ACA GCG CAT TAT AAG TGT CAA AAT CTC ATA GAA GTT GGA GTC AAA				1014
Phe Thr Ala His Tyr Lys Cys Gln Asn Leu Ile Glu Val Gly Val Lys	315	320	325	
ACC GAT AAA GGC TCC TGG ACT TTC AAC TTT AAC AGA TAAATCAGGC AAATAT				1066
Thr Asp Lys Gly Ser Trp Thr Phe Asn Phe Asn Arg	330	335		
GGACAATAGC ACAGACAGAG CAAAAATCCT TATAGAAGAG CTTAAAT				1114

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 338 amino acids

-250-

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

```

Met Phe Phe Lys Thr Tyr Gln Lys Leu Leu Gly Ala Ser Cys Leu Thr
 1      5      10      15
Leu Tyr Leu Ala Gly Cys Gly Ser Asp Ser Ser Glu Pro Leu Val Gly
 20      25      30
Ile Glu Lys Asn Ser Phe Asn Ser Thr Val Lys Ile Ile Ser Lys Thr
 35      40      45
Asp Asn Ile Glu Ile Gln Asp Leu Lys Leu Asn Arg Gly Asn Cys Glu
 50      55      60
His Asp Gln Asn Phe Leu Val Lys Leu Ile Gln Glu Thr Ala Asn Thr
 65      70      75      80
Tyr Leu Phe Ala Ser Glu Lys Glu Lys Ala Ile Lys Asn His Gln Ala
 85      90      95
Lys Ile Ala Arg Leu Gln Lys Asp Leu Glu Glu Leu Thr Gln His Val
 100     105     110
Gln Gln Ser Asn Asn Leu Asp Lys Leu Leu Glu Asn Gly Gly Leu Phe
 115     120     125
Val Ser Gly His Asp Tyr Lys Tyr Thr Lys Asp Asp Asn Pro Ile Tyr
 130     135     140
Val Val Lys Arg Met Leu Asp Asn Leu Asp Ser Tyr Lys Tyr Glu Ser
 145     150     155     160
Asp Asp Val Leu Asp Val Pro Tyr Glu Lys Leu Leu Glu Ile Ser Ile
 165     170     175
Ala Ile Glu Asp Thr Lys Asn Pro Lys Asp Tyr Pro Tyr Ile Asn Leu
 180     185     190
Lys Glu Leu Lys Lys Leu Ile Asp Ser Ile Ile Asp Asp His Gly Tyr
 195     200     205
Met Ala Asp Gly Phe Leu Asn Glu Tyr Ser Asn Arg Val Ser Lys Lys
 210     215     220
Gly Leu Gln Ile Leu Ala Lys Leu Lys Ser Met Trp Pro Ser Val Gly
 225     230     235     240
Lys Phe Tyr Phe Ala Ser Leu Lys Glu Ala Ile Pro Arg His Ala Lys
 245     250     255
Glu Val Thr Asp Lys Met Ile Ser Ser Glu Glu Lys Ser Ile Lys Ala
 260     265     270
Asn Gln Val Lys Leu Thr Glu Ala Lys Gln Asp Ile Asp Lys Met Glu
 275     280     285
Lys Ile Ile Lys Asp Leu Glu Ser Lys Lys Asn Thr Leu Ser Val Tyr
 290     295     300
Leu Lys Phe Gly Glu Ser Phe Thr Ala His Tyr Lys Cys Gln Asn Leu
 305     310     315     320
Ile Glu Val Gly Val Lys Thr Asp Lys Gly Ser Trp Thr Phe Asn Phe
 325     330     335
Asn Arg

```

(2) INFORMATION FOR SEQ ID NO:121:

-251-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 40...1026
- (D) OTHER INFORMATION:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 40...99
- (D) OTHER INFORMATION:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 100...1026
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

GGTTATACCG AAAAAACAAT ATGAAATCAA GGAGTTTGT ATG CAA CAG CGT CAT	54
Met Gln Gln Arg His	
-20	
TTA GGC CCT TTA AAA GTG GGT GCA TTA GCT CTA GGG TGC ATG GGC ATG	102
Leu Gly Pro Leu Lys Val Gly Ala Leu Ala Leu Gly Cys Met Gly Met	
-15 -10 -5 1	
ACT TAT GGG TAT GGG GAA GTC CAT GAT AAA AAG CAG ATG GTT AAA CTT	150
Thr Tyr Gly Tyr Gly Glu Val His Asp Lys Lys Gln Met Val Lys Leu	
5 10 15	
ATC CAT AAG GCT TTG GAA TTG GGT ATT AAC TTT TTT GAC ACT GCA GAG	198
Ile His Lys Ala Leu Glu Leu Gly Ile Asn Phe Phe Asp Thr Ala Glu	
20 25 30	
GCT TAT GGG GAA GAT AAT GAA AAG CTT TTA GCG AAG CGA TCA AGC CTT	246
Ala Tyr Gly Glu Asp Asn Glu Lys Leu Leu Ala Lys Arg Ser Ser Leu	
35 40 45	
ATT AAA GAC AAG GTT GTG GTA GCG AGC AAG TTT GGG ATT TAC TAC GCA	294
Ile Lys Asp Lys Val Val Val Ala Ser Lys Phe Gly Ile Tyr Tyr Ala	
50 55 60 65	
GAT CCT AAT GAC AAA TAC GCA ACC ATG TTT TTA GAC TCC AGT TCT AAC	342
Asp Pro Asn Asp Lys Tyr Ala Thr Met Phe Leu Asp Ser Ser Ser Asn	
70 75 80	
CGC ATT AAG AGT GCC ATT GAA GGG AGT TTG AAA CGC TTA AAA GTA GAA	390

-252-

Arg	Ile	Lys	Ser	Ala	Ile	Glu	Gly	Ser	Leu	Lys	Arg	Leu	Lys	Val	Glu		
			85					90					95				
TGC	ATT	GAT	TTA	TAC	TAC	CAA	CAC	CGC	ATG	GAT	ACT	AAC	ACG	CCC	ATA	438	
Cys	Ile	Asp	Leu	Tyr	Tyr	Gln	His	Arg	Met	Asp	Thr	Asn	Thr	Pro	Ile		
		100					105					110					
GAA	GAA	GTG	GCA	GAA	GTT	ATG	CAA	GCT	CTT	ATT	AAA	GAA	GGA	AAA	ATT	486	
Glu	Glu	Val	Ala	Glu	Val	Met	Gln	Ala	Leu	Ile	Lys	Glu	Gly	Lys	Ile		
		115				120					125						
AAA	GCT	TGG	GGG	ATG	AGT	GAG	GCA	GGG	TTA	TCT	AGC	ATC	CAA	AAA	GCC	534	
Lys	Ala	Trp	Gly	Met	Ser	Glu	Ala	Gly	Leu	Ser	Ser	Ile	Gln	Lys	Ala		
130					135					140					145		
CAT	CAA	ATT	TGC	CCT	TTA	AGC	GCG	TTG	CAG	AGC	GAA	TAT	TCC	TTG	TGG	582	
His	Gln	Ile	Cys	Pro	Leu	Ser	Ala	Leu	Gln	Ser	Glu	Tyr	Ser	Leu	Trp		
				150					155					160			
TGG	CGC	GAA	CCT	GAA	AAA	GAG	ATT	TTA	GGT	TTT	TTA	GAA	AAA	GAA	AAA	630	
Trp	Arg	Glu	Pro	Glu	Lys	Glu	Ile	Leu	Gly	Phe	Leu	Glu	Lys	Glu	Lys		
			165					170					175				
ATT	GGC	TTT	GTC	GCT	TTT	TCG	CCT	TTG	GGT	AAG	GGG	TTT	TTA	GGC	GCG	678	
Ile	Gly	Phe	Val	Ala	Phe	Ser	Pro	Leu	Gly	Lys	Gly	Phe	Leu	Gly	Ala		
		180					185					190					
AAA	TTT	GAA	AAA	AAT	GCT	ACC	TTC	GCT	AGT	GAA	GAT	TTT	AGA	AGC	GTT	726	
Lys	Phe	Glu	Lys	Asn	Ala	Thr	Phe	Ala	Ser	Glu	Asp	Phe	Arg	Ser	Val		
		195				200					205						
TCT	CCT	AGG	TTT	AAT	CAA	GAA	AAT	CTA	GCC	AAA	AAT	TAC	GTC	TTG	GTG	774	
Ser	Pro	Arg	Phe	Asn	Gln	Glu	Asn	Leu	Ala	Lys	Asn	Tyr	Val	Leu	Val		
210					215					220					225		
GAA	TTA	ATC	CAA	GAT	CAT	GCA	CAC	GCT	AAA	GGC	GTT	ACA	CCA	GCC	CAA	822	
Glu	Leu	Ile	Gln	Asp	His	Ala	His	Ala	Lys	Gly	Val	Thr	Pro	Ala	Gln		
				230					235					240			
CTG	GCT	CTC	TCG	TGG	ATT	TTG	CAC	ACG	CAA	AAA	ATC	ATT	GTC	CCT	CTC	870	
Leu	Ala	Leu	Ser	Trp	Ile	Leu	His	Thr	Gln	Lys	Ile	Ile	Val	Pro	Leu		
			245					250					255				
TTT	GGC	ACC	ACC	AAA	GAA	TCC	AGG	CTC	ATA	GAA	AAT	ATA	GGG	GCT	TTG	918	
Phe	Gly	Thr	Thr	Lys	Glu	Ser	Arg	Leu	Ile	Glu	Asn	Ile	Gly	Ala	Leu		
		260				265					270						
CAG	GTT	TCT	TGG	AGT	CAA	AAA	GAA	TTG	GAG	ATT	TTT	CAA	AAA	GAA	TTG	966	
Gln	Val	Ser	Trp	Ser	Gln	Lys	Glu	Leu	Glu	Ile	Phe	Gln	Lys	Glu	Leu		
		275				280					285						
ACT	GCA	ATC	AAA	ATA	GAA	GGG	GCC	CGC	TAC	CCT	GAA	AGA	ATC	AAT	GAA	1014	
Thr	Ala	Ile	Lys	Ile	Glu	Gly	Ala	Arg	Tyr	Pro	Glu	Arg	Ile	Asn	Glu		
		290			295					300					305		

-253-

ATG GTG AAT CAA TAAAAGTATT GGGTATTTAT AATTGCATTG GCTCTTTTAA AAGAG 1071
Met Val Asn Gln

ATTGAGCGTT ATTCCTGTG TGTCAGTGTG

1101

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Met	Gln	Gln	Arg	His	Leu	Gly	Pro	Leu	Lys	Val	Gly	Ala	Leu	Ala	Leu	-20	-15	-10	-5
Gly	Cys	Met	Gly	Met	Thr	Tyr	Gly	Tyr	Gly	Glu	Val	His	Asp	Lys	Lys				
				1				5					10						
Gln	Met	Val	Lys	Leu	Ile	His	Lys	Ala	Leu	Glu	Leu	Gly	Ile	Asn	Phe	15	20	25	
Phe	Asp	Thr	Ala	Glu	Ala	Tyr	Gly	Glu	Asp	Asn	Glu	Lys	Leu	Leu	Ala	30	35	40	
Lys	Arg	Ser	Ser	Leu	Ile	Lys	Asp	Lys	Val	Val	Val	Ala	Ser	Lys	Phe	45	50	55	60
Gly	Ile	Tyr	Tyr	Ala	Asp	Pro	Asn	Asp	Lys	Tyr	Ala	Thr	Met	Phe	Leu		65	70	75
Asp	Ser	Ser	Ser	Asn	Arg	Ile	Lys	Ser	Ala	Ile	Glu	Gly	Ser	Leu	Lys		80	85	90
Arg	Leu	Lys	Val	Glu	Cys	Ile	Asp	Leu	Tyr	Tyr	Gln	His	Arg	Met	Asp	95	100	105	
Thr	Asn	Thr	Pro	Ile	Glu	Glu	Val	Ala	Glu	Val	Met	Gln	Ala	Leu	Ile	110	115	120	
Lys	Glu	Gly	Lys	Ile	Lys	Ala	Trp	Gly	Met	Ser	Glu	Ala	Gly	Leu	Ser	125	130	135	140
Ser	Ile	Gln	Lys	Ala	His	Gln	Ile	Cys	Pro	Leu	Ser	Ala	Leu	Gln	Ser		145	150	155
Glu	Tyr	Ser	Leu	Trp	Trp	Arg	Glu	Pro	Glu	Lys	Glu	Ile	Leu	Gly	Phe		160	165	170
Leu	Glu	Lys	Glu	Lys	Ile	Gly	Phe	Val	Ala	Phe	Ser	Pro	Leu	Gly	Lys		175	180	185
Gly	Phe	Leu	Gly	Ala	Lys	Phe	Glu	Lys	Asn	Ala	Thr	Phe	Ala	Ser	Glu	190	195	200	
Asp	Phe	Arg	Ser	Val	Ser	Pro	Arg	Phe	Asn	Gln	Glu	Asn	Leu	Ala	Lys	205	210	215	220
Asn	Tyr	Val	Leu	Val	Glu	Leu	Ile	Gln	Asp	His	Ala	His	Ala	Lys	Gly		225	230	235
Val	Thr	Pro	Ala	Gln	Leu	Ala	Leu	Ser	Trp	Ile	Leu	His	Thr	Gln	Lys		240	245	250
Ile	Ile	Val	Pro	Leu	Phe	Gly	Thr	Thr	Lys	Glu	Ser	Arg	Leu	Ile	Glu				

[illegible]

(A) LENGTH: 955 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 126...806
(D) OTHER INFORMATION:

(A) NAME/KEY: Signal Sequence
(B) LOCATION: 126...237
(D) OTHER INFORMATION:

BNSDOCID: <WO 9821225A1_I_>

-255-

CTC ACC ACT CCT ATA GCC TTA CTC ATT GGC TTG TGG TTT GGG TTT CAA	410
Leu Thr Thr Pro Ile Ala Leu Leu Ile Gly Leu Trp Phe Gly Phe Gln	
45 50 55	
AAA CGC ATC GCT TTG GGG GTG TGG TTT TTC TTT AGC ATC TTA TTA GGT	458
Lys Arg Ile Ala Leu Gly Val Trp Phe Phe Phe Ser Ile Leu Leu Gly	
60 65 70	
GAA TTC ACC TTA AAA TCC CTT AAG CTT TTA GTG GCG CGC CCA CGG CCT	506
Glu Phe Thr Leu Lys Ser Leu Lys Leu Leu Val Ala Arg Pro Arg Pro	
75 80 85 90	
GTA ACC AAT GGC GAA TTG GTT TTC GCG CAT GGC TTT AGT TTC CCT AGC	554
Val Thr Asn Gly Glu Leu Val Phe Ala His Gly Phe Ser Phe Pro Ser	
95 100 105	
GGG CAT GCT TTG GCT TCA GCG CTT TTT TAC GGC TCT TTG GCG TTG TTG	602
Gly His Ala Leu Ala Ser Ala Leu Phe Tyr Gly Ser Leu Ala Leu Leu	
110 115 120	
TTA TGC TAT TCT AAC GCC AAC AAT CGC ATT AAA ACG ATT ATT GCT GTG	650
Leu Cys Tyr Ser Asn Ala Asn Asn Arg Ile Lys Thr Ile Ile Ala Val	
125 130 135	
GTT TTG CTT TTT TGG ATT TTT TTA ATG GCG TAT GAT AGG GTT TAT TTA	698
Val Leu Leu Phe Trp Ile Phe Leu Met Ala Tyr Asp Arg Val Tyr Leu	
140 145 150	
GGG GTG CAT TAC CCT AGC GAT GTT TTA GGA GGG TTT TTA TTA GGG ATT	746
Gly Val His Tyr Pro Ser Asp Val Leu Gly Gly Phe Leu Leu Gly Ile	
155 160 165 170	
GCT TGG TCG TGC TGC TCT TTA GCG CTT TAT TTA GGG TTT TTG AAA CGC	794
Ala Trp Ser Cys Cys Ser Leu Ala Leu Tyr Leu Gly Phe Leu Lys Arg	
175 180 185	
CCT TAT AAT CAA TAAAGGCTTT ATTTAACCAA ACGTACGACAA CTAAAATTTT TAAAA	851
Pro Tyr Asn Gln	
190	
TTCTATTTTTTGATAAAACT CATTCTCTTA AGGGGATAGG GGGTATTTTG CGATAATACC	911
CCCTTAACCC CCTTAAGAAA CCCCCTAACC CCCAAGACCG CTTT	955

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 227 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

-256-

- (A) NAME/KEY: Signal Sequence
 (B) LOCATION: 1...37
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

```

Met Val Phe Asp Arg Thr Ile Ser Val Arg Glu Lys Lys Ala Ala Lys
      -35                      -30                      -25
Thr Leu Gly Ile Ile Gly Ile Val Phe Phe Ile Leu Phe Gly Ile Val
      -20                      -15                      -10
Ile Ser Gly Val Ala Phe Gln Lys Glu Trp Val Gln Gln Leu Asp Leu
      -5                      1                      5                      10
Phe Phe Ile Asp Leu Ile Arg Asn Pro Ala Pro Ile Gln Lys Ser Ala
      15                      20                      25
Trp Leu Ser Phe Val Phe Phe Ser Thr Trp Phe Ala Gln Ser Lys Leu
      30                      35                      40
Thr Thr Pro Ile Ala Leu Leu Ile Gly Leu Trp Phe Gly Phe Gln Lys
      45                      50                      55
Arg Ile Ala Leu Gly Val Trp Phe Phe Phe Ser Ile Leu Leu Gly Glu
      60                      65                      70                      75
Phe Thr Leu Lys Ser Leu Lys Leu Leu Val Ala Arg Pro Arg Pro Val
      80                      85                      90
Thr Asn Gly Glu Leu Val Phe Ala His Gly Phe Ser Phe Pro Ser Gly
      95                      100                      105
His Ala Leu Ala Ser Ala Leu Phe Tyr Gly Ser Leu Ala Leu Leu Leu
      110                      115                      120
Cys Tyr Ser Asn Ala Asn Asn Arg Ile Lys Thr Ile Ile Ala Val Val
      125                      130                      135
Leu Leu Phe Trp Ile Phe Leu Met Ala Tyr Asp Arg Val Tyr Leu Gly
      140                      145                      150                      155
Val His Tyr Pro Ser Asp Val Leu Gly Gly Phe Leu Leu Gly Ile Ala
      160                      165                      170
Trp Ser Cys Cys Ser Leu Ala Leu Tyr Leu Gly Phe Leu Lys Arg Pro
      175                      180                      185
Tyr Asn Gln
      190

```

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1183 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 91...1032
 (D) OTHER INFORMATION:

-257-

- (A) NAME/KEY: Signal Sequence
 (B) LOCATION: 91...148
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

CTTAAAAGAA ACTTCGCAAA CCTTTTTATA TTATTTTAAA AGCACTAATA TTTATTATAT	60
TAGTTACAAC TATTATTGT AAAGGCTAAA ATG TTG AAA TTT AAA TAT GGT TTG	114
Met Leu Lys Phe Lys Tyr Gly Leu	
-15	
ATT TAT ATC GCG CTC ATA CTA GGA CTT CAA GCG ACA GAT TAT GAC AAT	162
Ile Tyr Ile Ala Leu Ile Leu Gly Leu Gln Ala Thr Asp Tyr Asp Asn	
-10 -5 1 5	
TTA GAA GAA GAA AAC CAA CAA TTA GAT GAA AAA ATA AAC CAT TTA AAG	210
Leu Glu Glu Glu Asn Gln Gln Leu Asp Glu Lys Ile Asn His Leu Lys	
10 15 20	
CAA CAG CTC ACC GAA AAA GGG GTT TCG CCC AAA GAG ATG GAT AAG GAT	258
Gln Gln Leu Thr Glu Lys Gly Val Ser Pro Lys Glu Met Asp Lys Asp	
25 30 35	
AAG TTT GAA GAA GAA TAC ATC AAT CGA TCT TAT CCT AAA ATT TCT TCC	306
Lys Phe Glu Glu Glu Tyr Ile Asn Arg Ser Tyr Pro Lys Ile Ser Ser	
40 45 50	
AAG AAA AAA GAG AAA TTG CTC AAA TCT TTT TCC ATA GCC GAT GAT AAG	354
Lys Lys Lys Glu Lys Leu Leu Lys Ser Phe Ser Ile Ala Asp Asp Lys	
55 60 65	
AGT GGG GTT TTT TTA GGG GGT GGG TAT GCT TAT GGG GAA CTT AAC TTG	402
Ser Gly Val Phe Leu Gly Gly Gly Tyr Ala Tyr Gly Glu Leu Asn Leu	
70 75 80 85	
TCT TAT CAA GGG GAA ATG TTA GAC AGA TAC GGC GCG AAT GCC CCT AGC	450
Ser Tyr Gln Gly Glu Met Leu Asp Arg Tyr Gly Ala Asn Ala Pro Ser	
90 95 100	
GCG TTT AAA AAC AAT ATC AAT ATT AAC GCT CCT GTT TCT ATG ATT AGC	498
Ala Phe Lys Asn Asn Ile Asn Ile Asn Ala Pro Val Ser Met Ile Ser	
105 110 115	
GCT AAA TTT GGG TAT CAA AAA TAC TTT GTG TCT TAT TTT GGG ACA CGA	546
Ala Lys Phe Gly Tyr Gln Lys Tyr Phe Val Ser Tyr Phe Gly Thr Arg	
120 125 130	
TTT TAT GGG GAT TTA TTG CTT GGG GGT GGG GCA TTA AAA GAG GAT GCA	594
Phe Tyr Gly Asp Leu Leu Leu Gly Gly Gly Ala Leu Lys Glu Asp Ala	
135 140 145	
ATC AAG CAG CCT GTA GGC TCG TTT ATT TAT GTT TTA GGG GCT GTC AAT	642
Ile Lys Gln Pro Val Gly Ser Phe Ile Tyr Val Leu Gly Ala Val Asn	

-258-

150	155	160	165	
ACC GAT TTA TTG TTT GAT ATG CCT TTA GAT TTT AAA ACT AAA AAG CAT				690
Thr Asp Leu Leu Phe Asp Met Pro Leu Asp Phe Lys Thr Lys Lys His				
	170	175	180	
TTT TTA GGC GTT TAT GCG GGT TTT GGG ATA GGG CTT ATG CTC TAT CAA				738
Phe Leu Gly Val Tyr Ala Gly Phe Gly Ile Gly Leu Met Leu Tyr Gln				
	185	190	195	
GAC AGG CCT AAT CAA AAC GGG AGG AAT TTA GTA GTG GGG GGC TAT TCA				786
Asp Arg Pro Asn Gln Asn Gly Arg Asn Leu Val Val Gly Gly Tyr Ser				
	200	205	210	
AGC CCT AAT TTT TTA TGG AAA TCT TTG ATT GAA GTG GAT TAC ACT TTT				834
Ser Pro Asn Phe Leu Trp Lys Ser Leu Ile Glu Val Asp Tyr Thr Phe				
	215	220	225	
AAT GTG GGC GTG AGT TTA ACG CTT TAT AGG AAA CAC CGT TTA GAG ATT				882
Asn Val Gly Val Ser Leu Thr Leu Tyr Arg Lys His Arg Leu Glu Ile				
	230	235	240	245
GGC ACA AAA TTG CCG ATT AGC TAT TTG AGA ATG GGA GTG GAA GAG GGA				930
Gly Thr Lys Leu Pro Ile Ser Tyr Leu Arg Met Gly Val Glu Glu Gly				
	250	255	260	
GCG ATT TAT CAA AAT AAA GAA GAT GAT GAG CGT TTG TTG GTT TCG GCT				978
Ala Ile Tyr Gln Asn Lys Glu Asp Asp Glu Arg Leu Leu Val Ser Ala				
	265	270	275	
AAC AAC CAG TTC AAG CGA TCC AGT TTT TTA TTA GTG AAT TAT GCG TTT				1026
Asn Asn Gln Phe Lys Arg Ser Ser Phe Leu Leu Val Asn Tyr Ala Phe				
	280	285	290	
ATT TTT TAAGGCTTGA TCTTGAGT AAGGTTTAAA ATTTTAGCGT TAGTCGTTTT AA				1084
Ile Phe				
	295			
TTTTAGGGGG TTATTTGATT TTTAACGCTT TAATCACAAA ACCCAGAGCT TTAAGTTTAA				1144
GTTTAAATAG CAAAGAGGGT GCGCTTAATG ACAATGATG				1183

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 314 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 1...19

-259-

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

Met Leu Lys Phe Lys Tyr Gly Leu Ile Tyr Ile Ala Leu Ile Leu Gly
              -15              -10              -5
Leu Gln Ala Thr Asp Tyr Asp Asn Leu Glu Glu Glu Asn Gln Gln Leu
              1              5              10
Asp Glu Lys Ile Asn His Leu Lys Gln Gln Leu Thr Glu Lys Gly Val
    15              20              25
Ser Pro Lys Glu Met Asp Lys Asp Lys Phe Glu Glu Glu Tyr Ile Asn
    30              35              40              45
Arg Ser Tyr Pro Lys Ile Ser Ser Lys Lys Lys Glu Lys Leu Leu Lys
              50              55              60
Ser Phe Ser Ile Ala Asp Asp Lys Ser Gly Val Phe Leu Gly Gly Gly
    65              70              75
Tyr Ala Tyr Gly Glu Leu Asn Leu Ser Tyr Gln Gly Glu Met Leu Asp
    80              85              90
Arg Tyr Gly Ala Asn Ala Pro Ser Ala Phe Lys Asn Asn Ile Asn Ile
    95              100              105
Asn Ala Pro Val Ser Met Ile Ser Ala Lys Phe Gly Tyr Gln Lys Tyr
   110              115              120              125
Phe Val Ser Tyr Phe Gly Thr Arg Phe Tyr Gly Asp Leu Leu Leu Gly
              130              135              140
Gly Gly Ala Leu Lys Glu Asp Ala Ile Lys Gln Pro Val Gly Ser Phe
              145              150              155
Ile Tyr Val Leu Gly Ala Val Asn Thr Asp Leu Leu Phe Asp Met Pro
              160              165              170
Leu Asp Phe Lys Thr Lys Lys His Phe Leu Gly Val Tyr Ala Gly Phe
              175              180              185
Gly Ile Gly Leu Met Leu Tyr Gln Asp Arg Pro Asn Gln Asn Gly Arg
   190              195              200              205
Asn Leu Val Val Gly Gly Tyr Ser Ser Pro Asn Phe Leu Trp Lys Ser
              210              215              220
Leu Ile Glu Val Asp Tyr Thr Phe Asn Val Gly Val Ser Leu Thr Leu
              225              230              235
Tyr Arg Lys His Arg Leu Glu Ile Gly Thr Lys Leu Pro Ile Ser Tyr
              240              245              250
Leu Arg Met Gly Val Glu Glu Gly Ala Ile Tyr Gln Asn Lys Glu Asp
              255              260              265
Asp Glu Arg Leu Leu Val Ser Ala Asn Asn Gln Phe Lys Arg Ser Ser
   270              275              280              285
Phe Leu Leu Val Asn Tyr Ala Phe Ile Phe
              290              295

```

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1851 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-260-

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 238...1665
 (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
 (B) LOCATION: 238...313
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

GAGCTAGTTT TAAAAAGTTA GTTTTGTTTT AAAAAGTTAA TACTATTTTG AAGCACTCCT	60
ATTCAGATGG CTAAGGCACA CAAGAAATTA GGGGACTCTG CTGTATTCCT ACCCTGAAGC	120
GTTACCCCTAA AATCCTATTG CATAGGTCTA AATAAGAGCT TAGGGATCAT TTTAGCCATA	180
AAAAGCTTAT GTTTTCATTA AAAATGTTAT GATACGCTCA AATAGTCAAG CAAAAAA ATG	240
	Met
	-25
TCA ATT AAA AGG GTT AGA TTG AAA ATA TTC GTT CTG TTG ATG TCG GTA	288
Ser Ile Lys Arg Val Arg Leu Lys Ile Phe Val Leu Leu Met Ser Val	
	-20 -15 -10
ATT TTA GGA ATA TCA TTA ACA GGT TGC ATA GGC TAT CGT ATG GAC TTA	336
Ile Leu Gly Ile Ser Leu Thr Gly Cys Ile Gly Tyr Arg Met Asp Leu	
	-5 1 5
GAA CAT TTT AAC ACG CTC TAT TAT GAA GAA AGC CCT AAA AAA GCT TAT	384
Glu His Phe Asn Thr Leu Tyr Tyr Glu Glu Ser Pro Lys Lys Ala Tyr	
	10 15 20
GAA TAT TCC AAA CAA TTC ACT AAG AAA AAA AAG AAC GCT CTT TTA TGG	432
Glu Tyr Ser Lys Gln Phe Thr Lys Lys Lys Lys Asn Ala Leu Leu Trp	
	25 30 35 40
GAC TTG CAA AAC GGC TTG AGC GCT TTA TAC GCC AGA GAT TAC CAG ACT	480
Asp Leu Gln Asn Gly Leu Ser Ala Leu Tyr Ala Arg Asp Tyr Gln Thr	
	45 50 55
TCT TTA GGG GTA TTA GAT CAA GCC GAG CAA CGC TTT GAT AAA ACG CAA	528
Ser Leu Gly Val Leu Asp Gln Ala Glu Gln Arg Phe Asp Lys Thr Gln	
	60 65 70
AGC GCT TTT ACA AGA GGG GCT GGT TAT GTG GGC GCT ACC ATG ATT AAT	576
Ser Ala Phe Thr Arg Gly Ala Gly Tyr Val Gly Ala Thr Met Ile Asn	
	75 80 85
GAT AAT GTG CGC GCT TAT GGG GGG AAT ATT TAT GAG GGC GTT TTA ATC	624
Asp Asn Val Arg Ala Tyr Gly Gly Asn Ile Tyr Glu Gly Val Leu Ile	
	90 95 100

-261-

AAT TAT TAC AAA GCG ATA GAC TAC ATG CTT TTA AAC GAT AGC GCG AAA Asn Tyr Tyr Lys Ala Ile Asp Tyr Met Leu Leu Asn Asp Ser Ala Lys 105 110 115 120	672
GCT AGG GTG CAA TTC AAC CGT GCG AAC GAA CGC CAG CGC AGG GCT AAA Ala Arg Val Gln Phe Asn Arg Ala Asn Glu Arg Gln Arg Arg Ala Lys 125 130 135	720
GAA TTT TAT TAT GAG GAA GTG CAA AAA GCC ATT AAA GAG ATC GAT TCT Glu Phe Tyr Tyr Glu Glu Val Gln Lys Ala Ile Lys Glu Ile Asp Ser 140 145 150	768
AGC AAA AAG CAC AAT ATT AAT ATG GAA CGC TCT AGG GTG GAA GTG AGC Ser Lys Lys His Asn Ile Asn Met Glu Arg Ser Arg Val Glu Val Ser 155 160 165	816
GAG ATT TTA AAC AAC ACC TAT TCT AAT TTA GAC AAA TAC GAA GCT TAT Glu Ile Leu Asn Asn Thr Tyr Ser Asn Leu Asp Lys Tyr Glu Ala Tyr 170 175 180	864
CAG GGC TTA CTT AAC CCG GCG GTT TCG TAT CTC TCA GGG TTG TTT TAC Gln Gly Leu Leu Asn Pro Ala Val Ser Tyr Leu Ser Gly Leu Phe Tyr 185 190 195 200	912
GCT TTA AAT GGG GAT GAG AAT AAG GGA TTA GGC TAT CTT AAT GAA GCC Ala Leu Asn Gly Asp Glu Asn Lys Gly Leu Gly Tyr Leu Asn Glu Ala 205 210 215	960
TAT GGG ATC AGT CAA AGC CCT TTT GTA GCC CAA GAC TTG GTT TTT TTC Tyr Gly Ile Ser Gln Ser Pro Phe Val Ala Gln Asp Leu Val Phe Phe 220 225 230	1008
AAA AAC CCT AAC AGG AGC CAT TTC ACT TGG ATC ATC ATT GAA GAT GGT Lys Asn Pro Asn Arg Ser His Phe Thr Trp Ile Ile Ile Glu Asp Gly 235 240 245	1056
AAA GAG CCG CAA AAA AGC GAA TTT AAA ATT GAT GTG CCT ATT TTT ATG Lys Glu Pro Gln Lys Ser Glu Phe Lys Ile Asp Val Pro Ile Phe Met 250 255 260	1104
ATC GAT TCG GTT TAT AAC GTG AGT ATA GCC TTG CCC AAG CTA GAA AAA Ile Asp Ser Val Tyr Asn Val Ser Ile Ala Leu Pro Lys Leu Glu Lys 265 270 275 280	1152
GGG GAA GCG TTT TAT CAA AAT TTC ACT CTC AAA GAT GGA GAA AAA GTA Gly Glu Ala Phe Tyr Gln Asn Phe Thr Leu Lys Asp Gly Glu Lys Val 285 290 295	1200
ACG CCC TTT GAC ACT TTA GCC TCA ATA GAT GCG GTG GTC GCT AGC GAA Thr Pro Phe Asp Thr Leu Ala Ser Ile Asp Ala Val Val Ala Ser Glu 300 305 310	1248
TTC AGG AAG CAG TTG CCC TAC ATT ATC ACT AGG GCT ATT TTA TCG GCC Phe Arg Lys Gln Leu Pro Tyr Ile Ile Thr Arg Ala Ile Leu Ser Ala 315 320 325	1296

-262-

ACT TTT AAA GTG GGC ATG CAA GCG GTG GCG AAC TAT TAT TTG GGG TTT	1344
Thr Phe Lys Val Gly Met Gln Ala Val Ala Asn Tyr Tyr Leu Gly Phe	
330 335 340	
GTT GGA GGG TTA GTA ACT TCC TTG TAT TCA GGT GTG AGC ACC TTT GCA	1392
Val Gly Gly Leu Val Thr Ser Leu Tyr Ser Gly Val Ser Thr Phe Ala	
345 350 355 360	
GAC ACT AGA AGC ACG AGC ATT TTT GCC CAT AAA ATC TAC CTC ATG CGC	1440
Asp Thr Arg Ser Thr Ser Ile Phe Ala His Lys Ile Tyr Leu Met Arg	
365 370 375	
ATT AAA AAC AAA GCC TTT GAA AGT TAT GAA GTT CGA GCC GAT TCC ATT	1488
Ile Lys Asn Lys Ala Phe Glu Ser Tyr Glu Val Arg Ala Asp Ser Ile	
380 385 390	
GAC GCT TTT TCG TTT TCA TTA AAG CCT TGT AAA AGA TCG CTT GAA AGC	1536
Asp Ala Phe Ser Phe Ser Leu Lys Pro Cys Lys Arg Ser Leu Glu Ser	
395 400 405	
CCT AAA ATC ATT GAC GCT AGG GAA TTG CTT TCT GGG TTT GTA GCA GCC	1584
Pro Lys Ile Ile Asp Ala Arg Glu Leu Leu Ser Gly Phe Val Ala Ala	
410 415 420	
CCA CAA ATC TTT TGC TCT AAC CGC CAT AAT ATT TTA TAC GTG CGC AGT	1632
Pro Gln Ile Phe Cys Ser Asn Arg His Asn Ile Leu Tyr Val Arg Ser	
425 430 435 440	
TTT AAA AAC GGG TTT GTT TTG AGT CGT TTA AAA TGATTTCAAA ACCCCCACCA	1685
Phe Lys Asn Gly Phe Val Leu Ser Arg Leu Lys	
445 450	
AAGGAATTTT AGTTTTTAAG TGTCGTTGGC ATTAAACGCA AACACGATAT AATTATAAAA	1745
CGATACGAAA ACCTAAATTA AGGGGAAGTC ATGGCTGATA GTTTAGCGGG CATTGATCAA	1805
GTTACGAGTT TGCATAAAAA TAACGAGTTA CAATTGTTGT GTTTCA	1851

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 476 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 1...25
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

-263-

Met	Ser	Ile	Lys	Arg	Val	Arg	Leu	Lys	Ile	Phe	Val	Leu	Leu	Met	Ser
-25					-20					-15					-10
Val	Ile	Leu	Gly	Ile	Ser	Leu	Thr	Gly	Cys	Ile	Gly	Tyr	Arg	Met	Asp
				-5					1				5		
Leu	Glu	His	Phe	Asn	Thr	Leu	Tyr	Tyr	Glu	Glu	Ser	Pro	Lys	Lys	Ala
		10					15					20			
Tyr	Glu	Tyr	Ser	Lys	Gln	Phe	Thr	Lys	Lys	Lys	Lys	Asn	Ala	Leu	Leu
	25					30						35			
Trp	Asp	Leu	Gln	Asn	Gly	Leu	Ser	Ala	Leu	Tyr	Ala	Arg	Asp	Tyr	Gln
40					45					50					55
Thr	Ser	Leu	Gly	Val	Leu	Asp	Gln	Ala	Glu	Gln	Arg	Phe	Asp	Lys	Thr
				60					65					70	
Gln	Ser	Ala	Phe	Thr	Arg	Gly	Ala	Gly	Tyr	Val	Gly	Ala	Thr	Met	Ile
			75					80					85		
Asn	Asp	Asn	Val	Arg	Ala	Tyr	Gly	Gly	Asn	Ile	Tyr	Glu	Gly	Val	Leu
		90					95					100			
Ile	Asn	Tyr	Tyr	Lys	Ala	Ile	Asp	Tyr	Met	Leu	Leu	Asn	Asp	Ser	Ala
	105					110						115			
Lys	Ala	Arg	Val	Gln	Phe	Asn	Arg	Ala	Asn	Glu	Arg	Gln	Arg	Arg	Ala
120					125					130					135
Lys	Glu	Phe	Tyr	Tyr	Glu	Glu	Val	Gln	Lys	Ala	Ile	Lys	Glu	Ile	Asp
				140					145					150	
Ser	Ser	Lys	Lys	His	Asn	Ile	Asn	Met	Glu	Arg	Ser	Arg	Val	Glu	Val
			155					160					165		
Ser	Glu	Ile	Leu	Asn	Asn	Thr	Tyr	Ser	Asn	Leu	Asp	Lys	Tyr	Glu	Ala
	170						175					180			
Tyr	Gln	Gly	Leu	Leu	Asn	Pro	Ala	Val	Ser	Tyr	Leu	Ser	Gly	Leu	Phe
	185					190					195				
Tyr	Ala	Leu	Asn	Gly	Asp	Glu	Asn	Lys	Gly	Leu	Gly	Tyr	Leu	Asn	Glu
200					205					210					215
Ala	Tyr	Gly	Ile	Ser	Gln	Ser	Pro	Phe	Val	Ala	Gln	Asp	Leu	Val	Phe
				220					225					230	
Phe	Lys	Asn	Pro	Asn	Arg	Ser	His	Phe	Thr	Trp	Ile	Ile	Ile	Glu	Asp
			235					240					245		
Gly	Lys	Glu	Pro	Gln	Lys	Ser	Glu	Phe	Lys	Ile	Asp	Val	Pro	Ile	Phe
	250					255						260			
Met	Ile	Asp	Ser	Val	Tyr	Asn	Val	Ser	Ile	Ala	Leu	Pro	Lys	Leu	Glu
	265					270					275				
Lys	Gly	Glu	Ala	Phe	Tyr	Gln	Asn	Phe	Thr	Leu	Lys	Asp	Gly	Glu	Lys
280					285					290					295
Val	Thr	Pro	Phe	Asp	Thr	Leu	Ala	Ser	Ile	Asp	Ala	Val	Val	Ala	Ser
				300					305					310	
Glu	Phe	Arg	Lys	Gln	Leu	Pro	Tyr	Ile	Ile	Thr	Arg	Ala	Ile	Leu	Ser
			315					320					325		
Ala	Thr	Phe	Lys	Val	Gly	Met	Gln	Ala	Val	Ala	Asn	Tyr	Tyr	Leu	Gly
	330					335						340			
Phe	Val	Gly	Gly	Leu	Val	Thr	Ser	Leu	Tyr	Ser	Gly	Val	Ser	Thr	Phe
	345					350					355				
Ala	Asp	Thr	Arg	Ser	Thr	Ser	Ile	Phe	Ala	His	Lys	Ile	Tyr	Leu	Met
360					365					370					375
Arg	Ile	Lys	Asn	Lys	Ala	Phe	Glu	Ser	Tyr	Glu	Val	Arg	Ala	Asp	Ser
			380						385					390	
Ile	Asp	Ala	Phe	Ser	Phe	Ser	Leu	Lys	Pro	Cys	Lys	Arg	Ser	Leu	Glu
		395					400						405		
Ser	Pro	Lys	Ile	Ile	Asp	Ala	Arg	Glu	Leu	Leu	Ser	Gly	Phe	Val	Ala

-264-

	410		415		420
Ala	Pro	Gln	Ile	Phe	Cys
	425		430		435
Ser	Phe	Lys	Asn	Gly	Phe
440			445		450

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 435 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...432
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

ATG	TTA	GAA	AAA	TTG	ATT	GAA	AGA	GTG	TTG	TTT	GCC	ACT	CGT	TGG	TTG	48
Met	Leu	Glu	Lys	Leu	Ile	Glu	Arg	Val	Leu	Phe	Ala	Thr	Arg	Trp	Leu	
1			5					10						15		
CTA	GCC	CCT	TTA	TGT	ATT	GCC	ATG	TCG	TTA	GTG	CTG	GTG	GTT	TTA	GGC	96
Leu	Ala	Pro	Leu	Cys	Ile	Ala	Met	Ser	Leu	Val	Leu	Val	Val	Leu	Gly	
			20					25						30		
TAT	GTG	TTC	ATG	AAA	GAG	TTG	TGG	CAC	ATG	CTC	AGC	CAT	TTA	AAC	ACG	144
Tyr	Val	Phe	Met	Lys	Glu	Leu	Trp	His	Met	Leu	Ser	His	Leu	Asn	Thr	
		35					40					45				
ATC	AGC	GAA	ACG	GAT	TTG	GTT	TTA	TCA	GCC	TTA	GGA	TTA	GTG	GAT	TTG	192
Ile	Ser	Glu	Thr	Asp	Leu	Val	Leu	Ser	Ala	Leu	Gly	Leu	Val	Asp	Leu	
		50				55					60					
TTG	TTT	ATG	GCC	GGG	CTT	GTT	TTA	ATG	GTG	TTA	CTC	GCC	AGT	TAT	GAA	240
Leu	Phe	Met	Ala	Gly	Leu	Val	Leu	Met	Val	Leu	Leu	Ala	Ser	Tyr	Glu	
65					70				75					80		
AGC	TTT	GTT	TCT	AAA	TTA	GAC	AAG	GTG	GAT	GCC	AGT	GAA	ATC	ACT	TGG	288
Ser	Phe	Val	Ser	Lys	Leu	Asp	Lys	Val	Asp	Ala	Ser	Glu	Ile	Thr	Trp	
				85				90						95		
CTA	AAG	CAC	ACG	GAT	TTT	AAC	GCT	TTA	AAA	TTA	AAG	GTT	TCA	CTC	TCC	336
Leu	Lys	His	Thr	Asp	Phe	Asn	Ala	Leu	Lys	Leu	Lys	Val	Ser	Leu	Ser	
			100					105						110		
ATT	GTA	GCG	ATT	TCA	GCG	ATT	TTC	TTG	CTC	AAA	CGC	TAC	ATG	AGT	TTA	384
Ile	Val	Ala	Ile	Ser	Ala	Ile	Phe	Leu	Leu	Lys	Arg	Tyr	Met	Ser	Leu	

-265-

115 120 125

GAA AGA TGT TTT ATC CCA GCA TTC CCT AAG GAT ACG CCC CCT ATC GCA T 433
 Glu Arg Cys Phe Ile Pro Ala Phe Pro Lys Asp Thr Pro Pro Ile Ala
 130 135 140

AA 435

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Met Leu Glu Lys Leu Ile Glu Arg Val Leu Phe Ala Thr Arg Trp Leu
 1 5 10 15
 Leu Ala Pro Leu Cys Ile Ala Met Ser Leu Val Leu Val Val Leu Gly
 20 25 30
 Tyr Val Phe Met Lys Glu Leu Trp His Met Leu Ser His Leu Asn Thr
 35 40 45
 Ile Ser Glu Thr Asp Leu Val Leu Ser Ala Leu Gly Leu Val Asp Leu
 50 55 60
 Leu Phe Met Ala Gly Leu Val Leu Met Val Leu Leu Ala Ser Tyr Glu
 65 70 75 80
 Ser Phe Val Ser Lys Leu Asp Lys Val Asp Ala Ser Glu Ile Thr Trp
 85 90 95

Leu Lys His Thr Asp Phe Asn Ala Leu Lys Leu Lys Val Ser Leu Ser
 100 105 110
 Ile Val Ala Ile Ser Ala Ile Phe Leu Leu Lys Arg Tyr Met Ser Leu
 115 120 125
 Glu Arg Cys Phe Ile Pro Ala Phe Pro Lys Asp Thr Pro Pro Ile Ala
 130 135 140

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2234 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 213...2081
- (D) OTHER INFORMATION:

-266-

- (A) NAME/KEY: Signal Sequence
 (B) LOCATION: 213...273
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

ATCATAAAAT GTAAAAATAC TCAAAGCATC GCATCAAGCA ATATAGCGAT CTGAAAAGAG	60
GCTCACAAATT GAGCTAAAGC CCGCTTTTTA GGGATAAATA AAAAGCGTTT TCAAATTGCA	120
TGGGTAACTT TATGGGGCGA AGCGTTTCTA AATTTTGGTA TAATCGCTAG AAATTGTGAG	180
AAAGATTCTA TCTTGTTTGA GTGGGGTTTC GC ATG CGT TTA TTA TTG TGG TGG	233
Met Arg Leu Leu Leu Trp Trp	
-20 -15	
GTA TTG GTA TTA TCG CTC TTT TTA AAT CCT TTG AGA GCG GTT GAA GAG	281
Val Leu Val Leu Ser Leu Phe Leu Asn Pro Leu Arg Ala Val Glu Glu	
-10 -5 1	
CAT GAA ACA GAT GCG GTG GAT TTG TTT TTG ATT TTC AAT CAA ATC AAC	329
His Glu Thr Asp Ala Val Asp Leu Phe Leu Ile Phe Asn Gln Ile Asn	
5 10 15	
CAG CTC AAT CAA GTC ATT GAA ACT TAC AAA AAA AAC CCT GAA AGA AGC	377
Gln Leu Asn Gln Val Ile Glu Thr Tyr Lys Lys Asn Pro Glu Arg Ser	
20 25 30 35	
GCT GAA ATC TCT CTG TAT AAC ACC CAA AAG AAT GAC TTG ATT AAA AGT	425
Ala Glu Ile Ser Leu Tyr Asn Thr Gln Lys Asn Asp Leu Ile Lys Ser	
40 45 50	
TTG ACT TCT AAA GTG TTG AAT GAA AGG GAT AAG ATC GGG ATT GAT ATC	473
Leu Thr Ser Lys Val Leu Asn Glu Arg Asp Lys Ile Gly Ile Asp Ile	
55 60 65	
AAT CAA AAT TTA AAA GAG CAG GAA AAA ATC AAA AAG CGT TTG TCT AAA	521
Asn Gln Asn Leu Lys Glu Gln Glu Lys Ile Lys Lys Arg Leu Ser Lys	
70 75 80	
AGC ATT AAT GGC GAT GAT TTC TAC ACT TTC ATG AAA GAC AGA TTG TCT	569
Ser Ile Asn Gly Asp Asp Phe Tyr Thr Phe Met Lys Asp Arg Leu Ser	
85 90 95	
TTA GAT ATT TTG TTG ATA GAT GAA ATT TTG TAT CGT TTT ATA GAT AAA	617
Leu Asp Ile Leu Leu Ile Asp Glu Ile Leu Tyr Arg Phe Ile Asp Lys	
100 105 110 115	
ATC AGG AGC AGT ATT GAT ATT TTT AGC GAA CAA AAA GAT GTA GAA AGC	665
Ile Arg Ser Ser Ile Asp Ile Phe Ser Glu Gln Lys Asp Val Glu Ser	
120 125 130	
ATC AGC GAT GCT TTC CTT TTG CGT TTA GGG CAA TTC AAA CTC TAC ACT	713

-267-

Ile Ser Asp Ala Phe Leu Leu Arg Leu Gly Gln Phe Lys Leu Tyr Thr	
135 140 145	
TTC CCT AAA AAT TTA GGC AAT GTC AAA ATG CAT GAA TTA GAG CAG ATG	761
Phe Pro Lys Asn Leu Gly Asn Val Lys Met His Glu Leu Glu Gln Met	
150 155 160	
TTT AGC GAT TAT GAA TTG CGT TTG AAC ACT TAC ACC GAA GTC TTG CGT	809
Phe Ser Asp Tyr Glu Leu Arg Leu Asn Thr Tyr Thr Glu Val Leu Arg	
165 170 175	
TAC ATT AAA AAC CAC CCT AAA GAA GTG CTT CCT AAA AAC TTG ATC ATG	857
Tyr Ile Lys Asn His Pro Lys Glu Val Leu Pro Lys Asn Leu Ile Met	
180 185 190 195	
GAA GTG AAT ATG GAT TTT GTG TTA AAC AAA ATC AGC AAG GTT TTG CCT	905
Glu Val Asn Met Asp Phe Val Leu Asn Lys Ile Ser Lys Val Leu Pro	
200 205 210	
TTC ACA ACC CAT AGC TTG CAA GTG AGT AAA ATC GTG CTA GCT TTG ACG	953
Phe Thr Thr His Ser Leu Gln Val Ser Lys Ile Val Leu Ala Leu Thr	
215 220 225	
ATT TTA GCC TTA TTG CTG GGT TTA AGG AAG TTG ATC ACT TGG CTT TTA	1001
Ile Leu Ala Leu Leu Leu Gly Leu Arg Lys Leu Ile Thr Trp Leu Leu	
230 235 240	
GCC TTA TTG TTA GAT CGT ATT TTT GAA ATC ATG CAG CGC AAT AAA AAA	1049
Ala Leu Leu Leu Asp Arg Ile Phe Glu Ile Met Gln Arg Asn Lys Lys	
245 250 255	
ATG CAT GTC AAT GTG CAA AAG AGC ATT GTT TCG CCG GTT TCT GTC TTT	1097
Met His Val Asn Val Gln Lys Ser Ile Val Ser Pro Val Ser Val Phe	
260 265 270 275	
TTA GCC CTA TTT AGT TGC GAT GTG GCT TTA GAT ATT TTC TAC TAC CCT	1145
Leu Ala Leu Phe Ser Cys Asp Val Ala Leu Asp Ile Phe Tyr Tyr Pro	
280 285 290	
AAC GCA TCG CCC CCT AAA GTT TCT ATG TGG GTG GGC GCG GTG TAT ATC	1193
Asn Ala Ser Pro Pro Lys Val Ser Met Trp Val Gly Ala Val Tyr Ile	
295 300 305	
ATG CTT TTA GCA TGG TTA GTG ATA GCG CTT TTT AAA GGC TAT GGG GAA	1241
Met Leu Leu Ala Trp Leu Val Ile Ala Leu Phe Lys Gly Tyr Gly Glu	
310 315 320	
GCG TTA GTT ACG AAT ATG GCT ACC AAA AGC ACG CAC AAT TTT AGA AAA	1289
Ala Leu Val Thr Asn Met Ala Thr Lys Ser Thr His Asn Phe Arg Lys	
325 330 335	
GAA GTG ATC AAC TTG ATT TTA AAA GTC GTG TAT TTT TTG ATC TTT ATT	1337
Glu Val Ile Asn Leu Ile Leu Lys Val Val Tyr Phe Leu Ile Phe Ile	
340 345 350 355	

GTC GCG CTT TTA GGG GTT TTG AAA CAA CTA GGG TTT AAC GTT TCA GCC Val Ala Leu Leu Gly Val Leu Lys Gln Leu Gly Phe Asn Val Ser Ala 360 365 370	1385
ATC ATC GCT TCT TTA GGG ATT GGG GGG TTA GCG GTG GCT TTG GCG GTT Ile Ile Ala Ser Leu Gly Ile Gly Gly Leu Ala Val Ala Leu Ala Val 375 380 385	1433
AAA GAT GTG TTA GCG AAT TTT TTT GCT TCG GTC ATT TTA TTA TTA GAC Lys Asp Val Leu Ala Asn Phe Phe Ala Ser Val Ile Leu Leu Leu Asp 390 395 400	1481
AAT TCG TTT TCT CAA GGG GAT TGG ATC GTG TGC GGT GAA GTG GAG GGC Asn Ser Phe Ser Gln Gly Asp Trp Ile Val Cys Gly Glu Val Glu Gly 405 410 415	1529
ACG GTG GTG GAA ATG GGG TTA AGG CGC ACC ACG ATC AGA GCC TTT GAC Thr Val Val Glu Met Gly Leu Arg Arg Thr Thr Ile Arg Ala Phe Asp 420 425 430 435	1577
AAC GCT CTT TTG TCC GTG CCT AAT TCA GAA TTA GCC GGA AAA CCC ATC Asn Ala Leu Leu Ser Val Pro Asn Ser Glu Leu Ala Gly Lys Pro Ile 440 445 450	1625
AGG AAT TGG AGC CGT CGT AAA GTG GGA AGG CGT ATT AAA ATG GAA ATA Arg Asn Trp Ser Arg Arg Lys Val Gly Arg Arg Ile Lys Met Glu Ile 455 460 465	1673
GGC TTA ACT TAT AGC TCC AGT CAA AGC GCT TTA CAG CTT TGC GTG AAA Gly Leu Thr Tyr Ser Ser Ser Gln Ser Ala Leu Gln Leu Cys Val Lys 470 475 480	1721
GAC ATT AAA GAA ATG TTA GAA AAC CAC CCT AAA ATC GCT AAC GGA GCC Asp Ile Lys Glu Met Leu Glu Asn His Pro Lys Ile Ala Asn Gly Ala 485 490 495	1769
GAT AGC GCT TTG CAA AAT GTG AGC GAT TAC CGC TAC ATG TTT AAA AAA Asp Ser Ala Leu Gln Asn Val Ser Asp Tyr Arg Tyr Met Phe Lys Lys 500 505 510 515	1817
GAT ATT GTT TCT ATT GAT GAT TTT TTA GGG TAT AAA AAC AAT TTG TTT Asp Ile Val Ser Ile Asp Asp Phe Leu Gly Tyr Lys Asn Asn Leu Phe 520 525 530	1865
GTC TTT TTA GAT CAG TTT GCG GAC AGC TCT ATT AAT ATT TTA GTG TAT Val Phe Leu Asp Gln Phe Ala Asp Ser Ser Ile Asn Ile Leu Val Tyr 535 540 545	1913
TGC TTT TCT AAG ACA GTG GTT TGG GAA GAG TGG CTA GAA GTC AAA GAA Cys Phe Ser Lys Thr Val Val Trp Glu Glu Trp Leu Glu Val Lys Glu 550 555 560	1961
GAT GTG ATG CTA AAA ATC ATG GGG ATT GTA GAA AAG CAC CAT TTG AGT Asp Val Met Leu Lys Ile Met Gly Ile Val Glu Lys His His Leu Ser 565 570 575	2009

-269-

```

TTT GCT TTC CCA TCA CAG AGT TTG TAT GTG GAG AGT TTG CCA GAA GTT      2057
Phe Ala Phe Pro Ser Gln Ser Leu Tyr Val Glu Ser Leu Pro Glu Val
580                      585                      590                      595

AGC CTG AAA GAA GGG GCT AAA ATC TGAAATTATT GGTAGATGTA TTCTTTGGTT      2111
Ser Leu Lys Glu Gly Ala Lys Ile
                      600

AAGGGGAAAG TGTTATCCAC GCTGTTGGTT AAAAGCAATT GGAATAAATC CGCGCTCCCC      2171
ACCCTAAAGG CGGATGCGCA AGTCCTTAA TACAGATCCC ACATGCGGAT AAAGCGTTCG      2231
TCA                                                                2234

```

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 623 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

```

Met Arg Leu Leu Leu Trp Trp Val Leu Val Leu Ser Leu Phe Leu Asn
-20                      -15                      -10                      -5
Pro Leu Arg Ala Val Glu Glu His Glu Thr Asp Ala Val Asp Leu Phe
                      1                      5                      10
Leu Ile Phe Asn Gln Ile Asn Gln Leu Asn Gln Val Ile Glu Thr Tyr
                      15                      20                      25
Lys Lys Asn Pro Glu Arg Ser Ala Glu Ile Ser Leu Tyr Asn Thr Gln
                      30                      35                      40
Lys Asn Asp Leu Ile Lys Ser Leu Thr Ser Lys Val Leu Asn Glu Arg
45                      50                      55                      60
Asp Lys Ile Gly Ile Asp Ile Asn Gln Asn Leu Lys Glu Gln Glu Lys
                      65                      70                      75
Ile Lys Lys Arg Leu Ser Lys Ser Ile Asn Gly Asp Asp Phe Tyr Thr
                      80                      85                      90
Phe Met Lys Asp Arg Leu Ser Leu Asp Ile Leu Leu Ile Asp Glu Ile
                      95                      100                      105
Leu Tyr Arg Phe Ile Asp Lys Ile Arg Ser Ser Ile Asp Ile Phe Ser
                      110                      115                      120
Glu Gln Lys Asp Val Glu Ser Ile Ser Asp Ala Phe Leu Leu Arg Leu
125                      130                      135                      140
Gly Gln Phe Lys Leu Tyr Thr Phe Pro Lys Asn Leu Gly Asn Val Lys
                      145                      150                      155
Met His Glu Leu Glu Gln Met Phe Ser Asp Tyr Glu Leu Arg Leu Asn

```

160							165							170						
Thr	Tyr	Thr	Glu	Val	Leu	Arg	Tyr	Ile	Lys	Asn	His	Pro	Lys	Glu	Val					
		175					180					185								
Leu	Pro	Lys	Asn	Leu	Ile	Met	Glu	Val	Asn	Met	Asp	Phe	Val	Leu	Asn					
		190				195					200									
Lys	Ile	Ser	Lys	Val	Leu	Pro	Phe	Thr	Thr	His	Ser	Leu	Gln	Val	Ser					
205					210					215					220					
Lys	Ile	Val	Leu	Ala	Leu	Thr	Ile	Leu	Ala	Leu	Leu	Leu	Gly	Leu	Arg					
				225					230					235						
Lys	Leu	Ile	Thr	Trp	Leu	Leu	Ala	Leu	Leu	Leu	Asp	Arg	Ile	Phe	Glu					
			240					245					250							
Ile	Met	Gln	Arg	Asn	Lys	Lys	Met	His	Val	Asn	Val	Gln	Lys	Ser	Ile					
		255					260					265								
Val	Ser	Pro	Val	Ser	Val	Phe	Leu	Ala	Leu	Phe	Ser	Cys	Asp	Val	Ala					
		270				275					280									
Leu	Asp	Ile	Phe	Tyr	Tyr	Pro	Asn	Ala	Ser	Pro	Pro	Lys	Val	Ser	Met					
285					290					295					300					
Trp	Val	Gly	Ala	Val	Tyr	Ile	Met	Leu	Leu	Ala	Trp	Leu	Val	Ile	Ala					
				305						310					315					
Leu	Phe	Lys	Gly	Tyr	Gly	Glu	Ala	Leu	Val	Thr	Asn	Met	Ala	Thr	Lys					
			320					325					330							
Ser	Thr	His	Asn	Phe	Arg	Lys	Glu	Val	Ile	Asn	Leu	Ile	Leu	Lys	Val					
		335					340					345								
Val	Tyr	Phe	Leu	Ile	Phe	Ile	Val	Ala	Leu	Leu	Gly	Val	Leu	Lys	Gln					
		350				355					360									
Leu	Gly	Phe	Asn	Val	Ser	Ala	Ile	Ile	Ala	Ser	Leu	Gly	Ile	Gly	Gly					
365					370					375					380					
Leu	Ala	Val	Ala	Leu	Ala	Val	Lys	Asp	Val	Leu	Ala	Asn	Phe	Phe	Ala					
				385					390						395					
Ser	Val	Ile	Leu	Leu	Leu	Asp	Asn	Ser	Phe	Ser	Gln	Gly	Asp	Trp	Ile					
			400					405					410							
Val	Cys	Gly	Glu	Val	Glu	Gly	Thr	Val	Val	Glu	Met	Gly	Leu	Arg	Arg					
		415					420					425								
Thr	Thr	Ile	Arg	Ala	Phe	Asp	Asn	Ala	Leu	Leu	Ser	Val	Pro	Asn	Ser					
		430				435						440								
Glu	Leu	Ala	Gly	Lys	Pro	Ile	Arg	Asn	Trp	Ser	Arg	Arg	Lys	Val	Gly					
445					450					455					460					
Arg	Arg	Ile	Lys	Met	Glu	Ile	Gly	Leu	Thr	Tyr	Ser	Ser	Ser	Gln	Ser					
				465					470					475						
Ala	Leu	Gln	Leu	Cys	Val	Lys	Asp	Ile	Lys	Glu	Met	Leu	Glu	Asn	His					
			480					485					490							
Pro	Lys	Ile</																		

-271-

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 432 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...429
- (D) OTHER INFORMATION:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 1...93
- (D) OTHER INFORMATION:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 94...429
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

ATG AAA AAA TTT TTT TCT CAA TCT TTA TTA GCT TTG ATT GTG TCT ATG	48
Met Lys Lys Phe Phe Ser Gln Ser Leu Leu Ala Leu Ile Val Ser Met	
-31 -30 -25 -20	
AAC GCG CTA CTG GCC ATG GAT GGC AAT GGC GTT TTT TTA GGG GCG GGT	96
Asn Ala Leu Leu Ala Met Asp Gly Asn Gly Val Phe Leu Gly Ala Gly	
-15 -10 -5 1	
TAT TTG CAA GGG CAA GCC CAA ATG CAT GCG GAT ATT AAT TCT CAA AAA	144
Tyr Leu Gln Gly Gln Ala Gln Met His Ala Asp Ile Asn Ser Gln Lys	
5 10 15	
CAA GCC ACT AAC GCT ACT ATC AAA GGC TTT GAT GCG CTT TTA GGG TAT	192
Gln Ala Thr Asn Ala Thr Ile Lys Gly Phe Asp Ala Leu Leu Gly Tyr	
20 25 30	
CAA TTT TTC TTT GGG AAA TAC TTT GGC TTG CGT GCT TAT GGG TTT TTT	240
Gln Phe Phe Phe Gly Lys Tyr Phe Gly Leu Arg Ala Tyr Gly Phe Phe	
35 40 45	
GAC TAC GCT CAT GCC AAT TCT ATT AGG CTT AAA AAC CCT AAC TAT AAC	288
Asp Tyr Ala His Ala Asn Ser Ile Arg Leu Lys Asn Pro Asn Tyr Asn	
50 55 60 65	
AGC GAA GTG GCG CAA TTG GCG GGT CAA ATT CTT GGG AAA CAA GAA ATC	336
Ser Glu Val Ala Gln Leu Ala Gly Gln Ile Leu Gly Lys Gln Glu Ile	

70

75

80

384

CTC ACT TAT GGG GGG GCT ATG GAT TTA ATG GTT AAT GTT CAT CAA TAA 432
Leu Thr Tyr Gly Gly Ala Met Asp Leu Met Val Asn Val His Gln
100 105 110

(D) TOPOLOGY: linear

(v) FRAGMENT TYPE: internal

Met	Lys	Lys	Phe	Phe	Ser	Gln	Ser	Leu	Leu	Ala	Leu	Ile	Val	Ser	Met
-31	-30					-25					-20				
Asn	Ala	Leu	Leu	Ala	Met	Asp	Gly	Asn	Gly	Val	Phe	Leu	Gly	Ala	Gly
-15					-10					-5					1
Tyr	Leu	Gln	Gly	Gln	Ala	Gln	Met	His	Ala	Asp	Ile	Asn	Ser	Gln	Lys
			5					10					15		
Gln	Ala	Thr	Asn	Ala	Thr	Ile	Lys	Gly	Phe	Asp	Ala	Leu	Leu	Gly	Tyr
		20					25					30			
Gln	Phe	Phe	Phe	Gly	Lys	Tyr	Phe	Gly	Leu	Arg	Ala	Tyr	Gly	Phe	Phe
	35					40				45					
Asp	Tyr	Ala	His	Ala	Asn	Ser	Ile	Arg	Leu	Lys	Asn	Pro	Asn	Tyr	Asn
50					55					60					65
Ser	Glu	Val	Ala	Gln	Leu	Ala	Gly	Gln	Ile	Leu	Gly	Lys	Gln	Glu	Ile
				70					75					80	
Asn	Arg	Leu	Thr	Ser	Leu	Ala	Asp	Pro	Lys	Thr	Phe	Glu	Pro	Asn	Met
			85					90					95		
Leu	Thr	Tyr	Gly	Gly	Ala	Met	Asp	Leu	Met	Val	Asn	Val	His	Gln	
		100					105					110			

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

- 273 -

(B) LOCATION: 1...333
(D) OTHER INFORMATION:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 1...60
(D) OTHER INFORMATION:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 61...333
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

ATG	AAA	ACC	TTT	AAA	AAC	CTG	CTC	TGT	TTT	AGC	CTG	ATC	GCT	ATG	AGT	48
Met	Lys	Thr	Phe	Lys	Asn	Leu	Leu	Cys	Phe	Ser	Leu	Ile	Ala	Met	Ser	
-20					-15					-10					-5	
TGG	CTC	CAA	GCG	GAC	ATG	TTG	GAT	AAT	TTC	ACT	AGG	GCC	ATT	AAC	AGC	96
Trp	Leu	Gln	Ala	Asp	Met	Leu	Asp	Asn	Phe	Thr	Arg	Ala	Ile	Asn	Ser	
				1				5					10			
TAC	ACC	ACT	AAA	AAG	CTT	AAT	GAA	ATC	AAG	GAT	CAA	GTC	AAT	AGC	GCT	144
Tyr	Thr	Thr	Lys	Lys	Leu	Asn	Glu	Ile	Lys	Asp	Gln	Val	Asn	Ser	Ala	
		15					20					25				
AAC	CCT	ACT	AAA	AAT	CAC	AAT	ACC	ACT	TAT	AAC	GCT	AAT	GGC	ATG	CTC	192
Asn	Pro	Thr	Lys	Asn	His	Asn	Thr	Thr	Tyr	Asn	Ala	Asn	Gly	Met	Leu	
	30					35					40					
ATT	AAC	ATT	GAT	TGT	AAA	GTC	TTA	AAA	AAT	AAC	TTC	TAT	TCG	GTG	TGT	240
Ile	Asn	Ile	Asp	Cys	Lys	Val	Leu	Lys	Asn	Asn	Phe	Tyr	Ser	Val	Cys	
45					50				55						60	
TAT	TCT	AGC	GAG	TTA	AAA	AAC	CCT	ATT	TAT	GGC	GTG	AGC	GTG	TTG	TTT	288
Tyr	Ser	Ser	Glu	Leu	Lys	Asn	Pro	Ile	Tyr	Gly	Val	Ser	Val	Leu	Phe	
				65				70						75		
GGG	GAT	TTA	GTG	GAT	AAA	AAT	AAT	ATT	GAA	AAA	CGC	TAT	GAG	TTT	TAA	336
Gly	Asp	Leu	Val	Asp	Lys	Asn	Asn	Ile	Glu	Lys	Arg	Tyr	Glu	Phe		
			80					85					90			

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 111 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

```
(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
```

-274-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Met	Lys	Thr	Phe	Lys	Asn	Leu	Leu	Cys	Phe	Ser	Leu	Ile	Ala	Met	Ser
-20					-15					-10					-5
Trp	Leu	Gln	Ala	Asp	Met	Leu	Asp	Asn	Phe	Thr	Arg	Ala	Ile	Asn	Ser
				1				5					10		
Tyr	Thr	Thr	Lys	Lys	Leu	Asn	Glu	Ile	Lys	Asp	Gln	Val	Asn	Ser	Ala
		15					20					25			
Asn	Pro	Thr	Lys	Asn	His	Asn	Thr	Thr	Tyr	Asn	Ala	Asn	Gly	Met	Leu
	30					35				40					
Ile	Asn	Ile	Asp	Cys	Lys	Val	Leu	Lys	Asn	Asn	Phe	Tyr	Ser	Val	Cys
45					50					55					60
Tyr	Ser	Ser	Glu	Leu	Lys	Asn	Pro	Ile	Tyr	Gly	Val	Ser	Val	Leu	Phe
				65					70					75	
Gly	Asp	Leu	Val	Asp	Lys	Asn	Asn	Ile	Glu	Lys	Arg	Tyr	Glu	Phe	
			80					85					90		

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2185 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 81...2069
- (D) OTHER INFORMATION:

- (A) NAME/KEY: Signal Sequence
- (B) LOCATION: 81...144
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

GTAAAAAATG	GCTTATCTGT	TCTAGCCTAC	TCCCCTTATT	TTTTCTTAAT	CCCTTAGCGG	60
CAGAAGATGA	TGGGTTTTTT	ATG GGG GTG AGT TAT CAA ACT TCT CTA GCT				110
		Met Gly Val Ser Tyr Gln Thr Ser Leu Ala				
		-20		-15		
ATT CAA AGG GTG GAT AAC TCA GGG CTT AAC GCC AGT CAA GCC GCA TCC						158
Ile Gln Arg Val Asp Asn Ser Gly Leu Asn Ala Ser Gln Ala Ala Ser						
-10		-5		1	5	
ACC TAC ATC CGC CAG AAC GCT ATC GCT CTA GAA TCT GCG GCG GTG CCT						206
Thr Tyr Ile Arg Gln Asn Ala Ile Ala Leu Glu Ser Ala Ala Val Pro						
	10		15		20	

-275-

TTA GCC TAT TAT TTA GAA GCG ATG GGC CAA CAA ACC AGG GTT TTA ATG	254
Leu Ala Tyr Tyr Leu Glu Ala Met Gly Gln Gln Thr Arg Val Leu Met	
25 30 35	
CAA ATG CTC TGC CCT GAT CCT TCC AAA CGC TGT TTG CTC TAT GCT GGA	302
Gln Met Leu Cys Pro Asp Pro Ser Lys Arg Cys Leu Leu Tyr Ala Gly	
40 45 50	
GGT TAT AAA AAC GGA TCA AGT AAT ACT AAC GGC GAT ACA GGC AAC AAC	350
Gly Tyr Lys Asn Gly Ser Ser Asn Thr Asn Gly Asp Thr Gly Asn Asn	
55 60 65	
CCC CCA AGA GGC AAT GTC AAT GCC ACC TTT GAT ATG CAA TCT CTA GTC	398
Pro Pro Arg Gly Asn Val Asn Ala Thr Phe Asp Met Gln Ser Leu Val	
70 75 80 85	
AAT AAT TTA AAC AAG CTC ACC CAA CTC ATC GGC GAG ACT TTA ATC CGT	446
Asn Asn Leu Asn Lys Leu Thr Gln Leu Ile Gly Glu Thr Leu Ile Arg	
90 95 100	
AAC CCT GAA AAT CTT TCT AAC GCC AAA GTC TTT AAT GTC AAA TTT GGC	494
Asn Pro Glu Asn Leu Ser Asn Ala Lys Val Phe Asn Val Lys Phe Gly	
105 110 115	
AAT CAA AGC ACT GTT ATT GCA TTG CCT GAG GGT CTA GCC AAT ACC ATG	542
Asn Gln Ser Thr Val Ile Ala Leu Pro Glu Gly Leu Ala Asn Thr Met	
120 125 130	
AAC GCT TTA AAC GAT GAT ATT ACC AAC GCT TTA ACC ACG CTC TGG TAT	590
Asn Ala Leu Asn Asp Asp Ile Thr Asn Ala Leu Thr Thr Leu Trp Tyr	
135 140 145	
AAC CAA ACC TTA ACG AAT AAA TCT TTT AAT AGC GGT AAT TCC GTG AAT	638
Asn Gln Thr Leu Thr Asn Lys Ser Phe Asn Ser Gly Asn Ser Val Asn	
150 155 160 165	
TTT AGC CCC CAA GTC TTG CAA CAC CTT TTA CAA GAC GGC TTA GCC ACA	686
Phe Ser Pro Gln Val Leu Gln His Leu Leu Gln Asp Gly Leu Ala Thr	
170 175 180	
AGT AAT CAA ACC ATT TGC AGC ACT CAA AAC CAA TGC ACC GCC ACC AAT	734
Ser Asn Gln Thr Ile Cys Ser Thr Gln Asn Gln Cys Thr Ala Thr Asn	
185 190 195	
GAA GCT AAA TCT ATC GCT CAA AAC GCC CAA AAC ATC TTC CAG GCT TTA	782
Glu Ala Lys Ser Ile Ala Gln Asn Ala Gln Asn Ile Phe Gln Ala Leu	
200 205 210	
ATG CAA GCA GGG ATT TTA GGG GGC TTA GCC AAT GAA AAG CAA TTT GGC	830
Met Gln Ala Gly Ile Leu Gly Gly Leu Ala Asn Glu Lys Gln Phe Gly	
215 220 225	
TTC ACT TAC AAC AAA GCC CCT AAT GGT AGC GAT TCC CAA CAA GGC TAC	878
Phe Thr Tyr Asn Lys Ala Pro Asn Gly Ser Asp Ser Gln Gln Gly Tyr	
230 235 240 245	

-276-

CAA AGC TTT AGC GGC CCG GGT TAT TAC ACT AAA AAC GGC GCT AAT GGC Gln Ser Phe Ser Gly Pro Gly Tyr Tyr Thr Lys Asn Gly Ala Asn Gly 250 255 260	926
ACT ACC CAA GCG CCC TTG AAA GCA TTA CCC GCT GGA GCG ACA ATT GGA Thr Thr Gln Ala Pro Leu Lys Ala Leu Pro Ala Gly Ala Thr Ile Gly 265 270 275	974
TCA GGC AAT GGC CAA TAC ACC TAC CAC CCC AGC TCG GCA GTC TAT TAT Ser Gly Asn Gly Gln Tyr Thr Tyr His Pro Ser Ser Ala Val Tyr Tyr 280 285 290	1022
TTA GCC GAT AGC ATC ATT GCT AAT GGC ATC ACC GCT TCT ATG ATT TTT Leu Ala Asp Ser Ile Ile Ala Asn Gly Ile Thr Ala Ser Met Ile Phe 295 300 305	1070
TCA GGC ATG CAA AAT TTC GCC AAT AAA GCC GCT AAA CTG ACA GGC ACT Ser Gly Met Gln Asn Phe Ala Asn Lys Ala Ala Lys Leu Thr Gly Thr 310 315 320 325	1118
TCA AGC TAT AGC CAG ATG CAA GAT GCG ATC AAT TAC GGG GAA AGC TTG Ser Ser Tyr Ser Gln Met Gln Asp Ala Ile Asn Tyr Gly Glu Ser Leu 330 335 340	1166
CTC AGT AAC ACC GTA GCG TAT GGG GAT TTC ATC ACC AAT TGG GTC GCC Leu Ser Asn Thr Val Ala Tyr Gly Asp Phe Ile Thr Asn Trp Val Ala 345 350 355	1214
CCC TAT TTG GAT TTA AAC AAC AAA GGT TTG AAT TTC TTG CCT AGC TAT Pro Tyr Leu Asp Leu Asn Asn Lys Gly Leu Asn Phe Leu Pro Ser Tyr 360 365 370	1262
GGG GGG CAA TTG AAT GGT GCT AAT CAT CAA ACC CCA CAA TTA ACC CCG Gly Gly Gln Leu Asn Gly Ala Asn His Gln Thr Pro Gln Leu Thr Pro 375 380 385	1310
CAA CAA GCC CAA CAA GAG CAA AAA GTC ATC ATG AAC CAA CTA GAG CAA Gln Gln Ala Gln Gln Glu Gln Lys Val Ile Met Asn Gln Leu Glu Gln 390 395 400 405	1358
GCC ACA AAC GCC CCC ACC CCC GCG CAA ATA AAC AGG ATT TTA GCC AAC Ala Thr Asn Ala Pro Thr Pro Ala Gln Ile Asn Arg Ile Leu Ala Asn 410 415 420	1406
CCC TAT TCC CCC ACG GCA AAA ACT TTA ATG GCT TAT GGG CTT TAT CGC Pro Tyr Ser Pro Thr Ala Lys Thr Leu Met Ala Tyr Gly Leu Tyr Arg 425 430 435	1454
TCT AAA GCA GTG ATT GGC GGG GTG ATT GAT GAA ATG CAA ACT AAA GTG Ser Lys Ala Val Ile Gly Gly Val Ile Asp Glu Met Gln Thr Lys Val 440 445 450	1502
AAT CAA GTC TAT CAA ATG GGC TTT GCT AGG AAT TTT TTG GAG CAT AAC Asn Gln Val Tyr Gln Met Gly Phe Ala Arg Asn Phe Leu Glu His Asn 455 460 465	1550

TCT Ser 470	AAT Asn	TCT Ser	AAT Asn	AAC Asn	ATG Met 475	AAC Asn	GGC Gly	TTT Phe	GGC Gly	GTG Val 480	AAA Lys	ATG Met	GGC Gly	TAT Tyr	AAG Lys 485	1598
CAA Gln	TTC Phe	TTT Phe	GGC Gly	AAA Lys 490	AAG Lys	CGC Arg	ATG Met	TTT Phe	GGG Gly 495	CTT Leu	AGG Arg	TAT Tyr	TAT Tyr	GGT Gly 500	TTT Phe	1646
TAT Tyr	GAT Asp	TTT Phe	GGT Gly 505	TAC Tyr	GCT Ala	CAA Gln	TTT Phe	GGC Gly 510	GCA Ala	GAA Glu	TCT Ser	TCT Ser	TTA Leu 515	GTG Val	AAA Lys	1694
GCC Ala	ACC Thr	CTC Leu 520	TCT Ser	AGC Ser	TAT Tyr	GGG Gly	GCA Ala 525	GGC Gly	ACA Thr	GAC Asp	TTT Phe	CTT Leu 530	TAT Tyr	AAT Asn	GTT Val	1742
TTT Phe 535	ACC Thr	CGA Arg	AAA Lys	AGA Arg	GGG Gly	ACT Thr 540	GAA Glu	GCG Ala	ATA Ile	GAT Asp	ATC Ile	GGT Gly	TTT Phe	TTT Phe	GCC Ala	1790
GGT Gly 550	ATC Ile	CAA Gln	CTT Leu	GCA Ala	GGG Gly 555	CAA Gln	ACT Thr	TGG Trp	AAA Lys	ACG Thr 560	AAT Asn	TTT Phe	TTA Leu	GAT Asp	CAA Gln 565	1838
GTG Val	GAT Asp	GGC Gly	AAC Asn	CAT His 570	CTT Leu	AAA Lys	CCC Pro	AAA Lys	GAC Asp	ACT Thr 575	TCT Ser	TTC Phe	CAA Gln	TTC Phe	CTT Leu 580	1886
TTT Phe	GAT Asp	TTA Leu	GGC Gly	ATA Ile	AGG Arg	ACC Thr	AAT Asn	TTT Phe 590	TCC Ser	AAA Lys	ATC Ile	GCT Ala	CAT His	CAA Gln	AAA Lys	1934
AGA Arg	TCC Ser	CGT Arg	TTT Phe	TCT Ser	CAA Gln	GGG Gly	ATA Ile 605	GAA Glu	TTT Phe	GGC Gly	CTT Leu	AAA Lys	ATA Ile	CCG Pro	GTG Val	1982
CTT Leu	TAT Tyr	CAC His	ACC Thr	TAT Tyr	TAC Tyr	CAA Gln 620	TCA Ser	GAA Glu	GGC Gly	GTT Val	ACA Thr	GCG Ala	AAG Lys	TAT Tyr	AGA Arg	2030
AGA Arg 630	GCC Ala	TTT Phe	AGT Ser	TTT Phe	TAT Tyr 635	GTG Val	GGC Gly	TAC Tyr	AAC Asn	ATA Ile 640	GGC Gly	TTT Phe	TGATTAAACA	AA		2081
ATAAGGGAAA	AATATGATAA		AAAAAGCTAG		AAAATTCATA		CCATTCTTTT		TAATTGGCTC						2141	
CCTCTTAGCT	GAAGACAATG		GCTGGTATAT		GTCTGTAGGC		TATC								2185	

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 663 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 278 -

```
(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(ix) FEATURE:
```

(A) NAME/KEY: Signal Sequence
(B) LOCATION: 1...21
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

Met	Gly	Val	Ser	Tyr	Gln	Thr	Ser	Leu	Ala	Ile	Gln	Arg	Val	Asp	Asn
-20						-15					-10				
Ser	Gly	Leu	Asn	Ala	Ser	Gln	Ala	Ala	Ser	Thr	Tyr	Ile	Arg	Gln	Asn
-5					1				5					10	
Ala	Ile	Ala	Leu	Glu	Ser	Ala	Ala	Val	Pro	Leu	Ala	Tyr	Tyr	Leu	Glu
			15					20					25		
Ala	Met	Gly	Gln	Gln	Thr	Arg	Val	Leu	Met	Gln	Met	Leu	Cys	Pro	Asp
			30				35					40			
Pro	Ser	Lys	Arg	Cys	Leu	Leu	Tyr	Ala	Gly	Gly	Tyr	Lys	Asn	Gly	Ser
	45					50					55				
Ser	Asn	Thr	Asn	Gly	Asp	Thr	Gly	Asn	Asn	Pro	Pro	Arg	Gly	Asn	Val
60					65					70					75
Asn	Ala	Thr	Phe	Asp	Met	Gln	Ser	Leu	Val	Asn	Asn	Leu	Asn	Lys	Leu
			80						85					90	
Thr	Gln	Leu	Ile	Gly	Glu	Thr	Leu	Ile	Arg	Asn	Pro	Glu	Asn	Leu	Ser
			95					100					105		
Asn	Ala	Lys	Val	Phe	Asn	Val	Lys	Phe	Gly	Asn	Gln	Ser	Thr	Val	Ile
		110					115					120			
Ala	Leu	Pro	Glu	Gly	Leu	Ala	Asn	Thr	Met	Asn	Ala	Leu	Asn	Asp	Asp
	125					130					135				
Ile	Thr	Asn	Ala	Leu	Thr	Thr	Leu	Trp	Tyr	Asn	Gln	Thr	Leu	Thr	Asn
140					145					150					155
Lys	Ser	Phe	Asn	Ser	Gly	Asn	Ser	Val	Asn	Phe	Ser	Pro	Gln	Val	Leu
			160						165					170	
Gln	His	Leu	Leu	Gln	Asp	Gly	Leu	Ala	Thr	Ser	Asn	Gln	Thr	Ile	Cys
			175					180					185		
Ser	Thr	Gln	Asn	Gln	Cys	Thr	Ala	Thr	Asn	Glu	Ala	Lys	Ser	Ile	Ala
		190				195					200				
Gln	Asn	Ala	Gln	Asn	Ile	Phe	Gln	Ala	Leu	Met	Gln	Ala	Gly	Ile	Leu
	205					210					215				
Gly	Gly	Leu	Ala	Asn	Glu	Lys	Gln	Phe	Gly	Phe	Thr	Tyr	Asn	Lys	Ala
220					225					230					235
Pro	Asn	Gly	Ser	Asp	Ser	Gln	Gln	Gly	Tyr	Gln	Ser	Phe	Ser	Gly	Pro
			240						245					250	
Gly	Tyr	Tyr	Thr	Lys	Asn	Gly	Ala	Asn	Gly	Thr	Thr	Gln	Ala	Pro	Leu
			255					260					265		
Lys	Ala	Leu	Pro	Ala	Gly	Ala	Thr	Ile	Gly	Ser	Gly	Asn	Gly	Gln	Tyr
		270				275						280			
Thr	Tyr	His	Pro	Ser	Ser	Ala	Val	Tyr	Tyr	Leu	Ala	Asp	Ser	Ile	Ile
	285					290					295				
Ala	Asn	Gly	Ile	Thr	Ala	Ser	Met	Ile	Phe	Ser	Gly	Met	Gln	Asn	Phe
300					305					310					315
Ala	Asn	Lys	Ala	Ala	Lys	Leu	Thr	Gly	Thr	Ser	Ser	Tyr	Ser	Gln	Met

-279-

320																325																330															
Gln	Asp	Ala	Ile	Asn	Tyr	Gly	Glu	Ser	Leu	Leu	Ser	Asn	Thr	Val	Ala																																
335																340																345															
Tyr	Gly	Asp	Phe	Ile	Thr	Asn	Trp	Val	Ala	Pro	Tyr	Leu	Asp	Leu	Asn																																
350																355																360															
Asn	Lys	Gly	Leu	Asn	Phe	Leu	Pro	Ser	Tyr	Gly	Gly	Gln	Leu	Asn	Gly																																
365																370																375															
Ala	Asn	His	Gln	Thr	Pro	Gln	Leu	Thr	Pro	Gln	Gln	Ala	Gln	Gln	Glu																																
380																385																390															
Gln	Lys	Val	Ile	Met	Asn	Gln	Leu	Glu	Gln	Ala	Thr	Asn	Ala	Pro	Thr																																
400																405																410															
Pro	Ala	Gln	Ile	Asn	Arg	Ile	Leu	Ala	Asn	Pro	Tyr	Ser	Pro	Thr	Ala																																
415																420																425															
Lys	Thr	Leu	Met	Ala	Tyr	Gly	Leu	Tyr	Arg	Ser	Lys	Ala	Val	Ile	Gly																																
430																435																440															
Gly	Val	Ile	Asp	Glu	Met	Gln	Thr	Lys	Val	Asn	Gln	Val	Tyr	Gln	Met																																
445																450																455															
Gly	Phe	Ala	Arg	Asn	Phe	Leu	Glu	His	Asn	Ser	Asn	Ser	Asn	Asn	Met																																
460																465																470															
Asn	Gly	Phe	Gly	Val	Lys	Met	Gly	Tyr	Lys	Gln	Phe	Phe	Gly	Lys	Lys																																
480																485																490															
Arg	Met	Phe	Gly	Leu	Arg	Tyr	Tyr	Gly	Phe	Tyr	Asp	Phe	Gly	Tyr	Ala																																
495																500																505															
Gln	Phe	Gly	Ala	Glu	Ser	Ser	Leu	Val	Lys	Ala	Thr	Leu	Ser	Ser	Tyr																																
510																515																520															
Gly	Ala	Gly	Thr	Asp	Phe	Leu	Tyr	Asn	Val	Phe	Thr	Arg	Lys	Arg	Gly																																
525																530																535															
Thr	Glu	Ala	Ile	Asp	Ile	Gly	Phe	Phe	Ala	Gly	Ile	Gln	Leu	Ala	Gly																																
540																545																550															
Gln	Thr	Trp	Lys	Thr	Asn	Phe	Leu	Asp	Gln	Val	Asp	Gly	Asn	His	Leu																																
560																565																570															
Lys	Pro	Lys	Asp	Thr	Ser	Phe	Gln	Phe	Leu	Phe	Asp	Leu	Gly	Ile	Arg																																
575																580																585															
Thr	Asn	Phe	Ser	Lys	Ile	Ala	His	Gln	Lys	Arg	Ser	Arg	Phe	Ser	Gln																																
590																595																600															
Gly	Ile	Glu	Phe	Gly	Leu	Lys	Ile	Pro	Val	Leu	Tyr	His	Thr	Tyr	Tyr																																
605																610																615															
Gln	Ser	Glu	Gly	Val	Thr	Ala	Lys	Tyr	Arg	Arg	Ala	Phe	Ser	Phe	Tyr																																
620																625																630															
Val	Gly	Tyr	Asn	Ile	Gly	Phe																																									
640																																															

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 51...1160

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

ATTATTTTT	TA	ATCTTGC	ATG	AAATCTT	AAA	TATAGA	AATTA	GTCCCTT	TGG	ATG	GGA	Met	Gly		56	
TTT	TCN	CTC	GCG	CTA	GGC	TAT	TTG	TGT	TTG	TTT	ATA	TTC	GTT	TTA	AGC	104
Phe	Xaa	Leu	Ala	Leu	Gly	Tyr	Leu	Cys	Leu	Phe	Ile	Phe	Val	Leu	Ser	
		5					10				15					
GCT	TCT	TTA	ATC	TCT	GAA	AAA	GCC	TTA	TCC	AAG	CAG	TAT	TTG	CAA	ACC	152
Ala	Ser	Leu	Ile	Ser	Glu	Lys	Ala	Leu	Ser	Lys	Gln	Tyr	Leu	Gln	Thr	
	20					25					30					
GCT	AAA	GAT	AAA	ATC	ACC	TCT	TTA	AAG	AAT	TTA	AAA	GTC	ATC	GCC	ATT	200
Ala	Lys	Asp	Lys	Ile	Thr	Ser	Leu	Lys	Asn	Leu	Lys	Val	Ile	Ala	Ile	
35					40					45					50	
ACC	GGA	AGC	TTT	GGG	AAA	ACC	AGC	ACC	AAA	AAT	TTC	TTG	CTT	CAA	ATC	248
Thr	Gly	Ser	Phe	Gly	Lys	Thr	Ser	Thr	Lys	Asn	Phe	Leu	Leu	Gln	Ile	
				55					60					65		
TTA	CAA	ACC	ACA	TTC	AAC	GCG	CAT	GCA	AGC	CCC	AAA	AGC	GTC	AAT	ACC	296
Leu	Gln	Thr	Thr	Phe	Asn	Ala	His	Ala	Ser	Pro	Lys	Ser	Val	Asn	Thr	
			70				75						80			
CTT	TTA	GGG	CTT	GCG	AAT	GAT	ATT	AAT	CAG	AAT	TTA	GAC	GAT	AGG	AGT	344
Leu	Leu	Gly	Leu	Ala	Asn	Asp	Ile	Asn	Gln	Asn	Leu	Asp	Asp	Arg	Ser	
		85					90					95				
GAA	ATC	TAT	ATC	GCT	GAA	GCC	GGG	GCA	AGG	AAT	AAG	GGC	GAT	ATT	AAA	392
Glu	Ile	Tyr	Ile	Ala	Glu	Ala	Gly	Ala	Arg	Asn	Lys	Gly	Asp	Ile	Lys	
	100					105					110					
GAA	ATC	ACC	TGT	CTC	ATT	GAA	CCG	CAC	CTT	GTT	GTG	GTT	GCA	GAA	GTG	440
Glu	Ile	Thr	Cys	Leu	Ile	Glu	Pro	His	Leu	Val	Val	Val	Ala	Glu	Val	
115					120					125					130	
GGC	GAA	CAG	CAT	TTA	GAA	TAC	TTT	AAA	ACT	TTA	GAA	AAT	ATT	TGC	GAG	488
Gly	Glu	Gln	His	Leu	Glu	Tyr	Phe	Lys	Thr	Leu	Glu	Asn	Ile	Cys	Glu	
			135					140						145		
ACT	AAA	GCG	GAA	TTA	TTG	GAT	TCC	AAA	CGC	TTA	GAA	AAA	GCC	TTT	TGT	536
Thr	Lys	Ala	Glu	Leu	Leu	Asp	Ser	Lys	Arg	Leu	Glu	Lys	Ala	Phe	Cys	
		150						155					160			
TAC	TCG	GTG	GAA	AAG	ATC	AAG	CCC	TAT	GCC	CCT	AAA	GAT	AGC	CCT	TTA	584
Tyr	Ser	Val	Glu	Lys	Ile	Lys	Pro	Tyr	Ala	Pro	Lys	Asp	Ser	Pro	Leu	
		165					170					175				
ATA	GAC	TAT	TCT	AGC	CTG	GTT	AAA	AAC	ATC	CAA	TCC	ACT	TTA	AAA	GGC	632

-281-

Ile	Asp	Tyr	Ser	Ser	Leu	Val	Lys	Asn	Ile	Gln	Ser	Thr	Leu	Lys	Gly		
180						185					190						
ACT	TCT	TTT	GAA	ATG	CTT	ATA	GGT	AGC	GTT	TGG	GAA	AGA	TTT	GAA	ACA	680	
Thr	Ser	Phe	Glu	Met	Leu	Ile	Gly	Ser	Val	Trp	Glu	Arg	Phe	Glu	Thr		
195					200					205					210		
AAG	GTT	CTA	GGG	GAG	TTT	AGC	GCT	TAT	AAT	ATC	GCT	TCA	GCC	ATT	TTA	728	
Lys	Val	Leu	Gly	Glu	Phe	Ser	Ala	Tyr	Asn	Ile	Ala	Ser	Ala	Ile	Leu		
				215					220					225			
ATC	GCT	AAG	CAT	TTA	GGC	TTA	GAG	ACC	GAA	AGG	ATC	AAA	CGG	CTT	GTT	776	
Ile	Ala	Lys	His	Leu	Gly	Leu	Glu	Thr	Glu	Arg	Ile	Lys	Arg	Leu	Val		
			230					235					240				
TTA	GAA	CTC	AAC	CCT	ATT	GCT	CAT	CGT	TTG	CAA	CTT	TTG	GAA	GTG	AAT	824	
Leu	Glu	Leu	Asn	Pro	Ile	Ala	His	Arg	Leu	Gln	Leu	Leu	Glu	Val	Asn		
		245					250					255					
CAA	AAA	ATC	ATC	ATA	GAC	GAT	AGC	TTT	AAT	GGG	AAT	TTA	AAG	GGC	ATG	872	
Gln	Lys	Ile	Ile	Ile	Asp	Asp	Ser	Phe	Asn	Gly	Asn	Leu	Lys	Gly	Met		
	260					265					270						
TTA	GAG	GGC	ATT	CGT	TTA	GCG	AGT	TTG	CAC	AAA	GGG	CGT	AAA	GTC	ATT	920	
Leu	Glu	Gly	Ile	Arg	Leu	Ala	Ser	Leu	His	Lys	Gly	Arg	Lys	Val	Ile		
275					280					285				290			
GTA	ACA	CCG	GGC	TTA	GTG	GAA	AGC	AAT	ACA	GAA	AGT	AAT	GAG	GCT	TTA	968	
Val	Thr	Pro	Gly	Leu	Val	Glu	Ser	Asn	Thr	Glu	Ser	Asn	Glu	Ala	Leu		
				295					300					305			
GCG	CAA	AAA	ATA	GAC	GGG	GTT	TTT	GAT	GTC	GCT	ATC	ATC	ACA	GGG	GAG	1016	
Ala	Gln	Lys	Ile	Asp	Gly	Val	Phe	Asp	Val	Ala	Ile	Ile	Thr	Gly	Glu		
			310					315					320				
TTG	AAT	TCC	AAA	ACG	ATT	GCT	TCA	CAA	TTG	AAA	ACC	CCC	CAA	AAA	ATC	1064	
Leu	Asn	Ser	Lys	Thr	Ile	Ala	Ser	Gln	Leu	Lys	Thr	Pro	Gln	Lys	Ile		
		325					330					335					
TTA	CTC	AAG	GAT	AAG	GCG	CAA	TTG	GAA	AAT	ATC	TTA	CAA	GCC	ACC	ACG	1112	
Leu	Leu	Lys	Asp	Lys	Ala	Gln	Leu	Glu	Asn	Ile	Leu	Gln	Ala	Thr	Thr		
		340				345					350						
ATT	CAA	GGC	GAT	TTG	ATT	TTA	TTC	GCT	AAT	GAC	GCC	CCT	AAT	TAC	ATT	T	1161
Ile	Gln	Gly	Asp	Leu	Ile	Leu	Phe	Ala	Asn	Asp	Ala	Pro	Asn	Tyr	Ile		
355					360				365					370			
AGGAAATGAA	CATGCAACAT	TTATACGCTC	CTTGCGCGCA	AAGTTATTTG	AA											1213	

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 370 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

-282-

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

```

Met Gly Phe Xaa Leu Ala Leu Gly Tyr Leu Cys Leu Phe Ile Phe Val
 1           5           10           15
Leu Ser Ala Ser Leu Ile Ser Glu Lys Ala Leu Ser Lys Gln Tyr Leu
          20           25           30
Gln Thr Ala Lys Asp Lys Ile Thr Ser Leu Lys Asn Leu Lys Val Ile
          35           40           45
Ala Ile Thr Gly Ser Phe Gly Lys Thr Ser Thr Lys Asn Phe Leu Leu
          50           55           60
Gln Ile Leu Gln Thr Thr Phe Asn Ala His Ala Ser Pro Lys Ser Val
          65           70           75           80
Asn Thr Leu Leu Gly Leu Ala Asn Asp Ile Asn Gln Asn Leu Asp Asp
          85           90           95
Arg Ser Glu Ile Tyr Ile Ala Glu Ala Gly Ala Arg Asn Lys Gly Asp
          100          105          110
Ile Lys Glu Ile Thr Cys Leu Ile Glu Pro His Leu Val Val Val Ala
          115          120          125
Glu Val Gly Glu Gln His Leu Glu Tyr Phe Lys Thr Leu Glu Asn Ile
          130          135          140
Cys Glu Thr Lys Ala Glu Leu Leu Asp Ser Lys Arg Leu Glu Lys Ala
          145          150          155          160
Phe Cys Tyr Ser Val Glu Lys Ile Lys Pro Tyr Ala Pro Lys Asp Ser
          165          170          175
Pro Leu Ile Asp Tyr Ser Ser Leu Val Lys Asn Ile Gln Ser Thr Leu
          180          185          190
Lys Gly Thr Ser Phe Glu Met Leu Ile Gly Ser Val Trp Glu Arg Phe
          195          200          205
Glu Thr Lys Val Leu Gly Glu Phe Ser Ala Tyr Asn Ile Ala Ser Ala
          210          215          220
Ile Leu Ile Ala Lys His Leu Gly Leu Glu Thr Glu Arg Ile Lys Arg
          225          230          235          240
Leu Val Leu Glu Leu Asn Pro Ile Ala His Arg Leu Gln Leu Leu Glu
          245          250          255
Val Asn Gln Lys Ile Ile Ile Asp Asp Ser Phe Asn Gly Asn Leu Lys
          260          265          270
Gly Met Leu Glu Gly Ile Arg Leu Ala Ser Leu His Lys Gly Arg Lys
          275          280          285
Val Ile Val Thr Pro Gly Leu Val Glu Ser Asn Thr Glu Ser Asn Glu
          290          295          300
Ala Leu Ala Gln Lys Ile Asp Gly Val Phe Asp Val Ala Ile Ile Thr
          305          310          315          320
Gly Glu Leu Asn Ser Lys Thr Ile Ala Ser Gln Leu Lys Thr Pro Gln
          325          330          335
Lys Ile Leu Leu Lys Asp Lys Ala Gln Leu Glu Asn Ile Leu Gln Ala
          340          345          350
Thr Thr Ile Gln Gly Asp Leu Ile Leu Phe Ala Asn Asp Ala Pro Asn
          355          360          365
Tyr Ile
          370

```

-283-

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 82...270
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

```

ACTTAAAGGC ATAAAAACCT TAAGCTTTTT GAGTTTCAAA AGGGTTTCAA GCTTTTTATA      60
AGACTTTTTT TGAATGAGTA A GGA GAA AAT ATT TTG TTC CAT AAA CTG ATC      111
                Gly Glu Asn Ile Leu Phe His Lys Leu Ile
                1                5                10

TTA ACA TGC TTT TTA GCG CTT GTA GCA ATA ACC ATT CAA GCT TGC GGT      159
Leu Thr Cys Phe Leu Ala Leu Val Ala Ile Thr Ile Gln Ala Cys Gly
                15                20                25

TAT AAA GCC CCT CCA TTC AAT GAA AAA CCC GCT AAA AAA ACT TCA AAC      207
Tyr Lys Ala Pro Pro Phe Asn Glu Lys Pro Ala Lys Lys Thr Ser Asn
                30                35                40

AGC TCT AAT TCT TCT ATG CAA ACG CCC ACC AAC AGC ACC ACG CCA GAA      255
Ser Ser Asn Ser Ser Met Gln Thr Pro Thr Asn Ser Thr Thr Pro Glu
                45                50                55

TTT TTA AAT CAG CCT TAAAATCACT GCTCTTGTTT AAGGGCTTTG ATTTCTAGGG T      311
Phe Leu Asn Gln Pro
                60

TTTTGTGGCT AACTTTTGAN STTCGCTTTC ATCATGCGTT ACCATAATG      360

```

(2) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 63 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

Gly Glu Asn Ile Leu Phe His Lys Leu Ile Leu Thr Cys Phe Leu Ala

-284-

1	5	10	15
Leu Val Ala	Ile Thr Ile Gln Ala Cys Gly Tyr Lys Ala	Pro Pro Phe	
20	25	30	
Asn Glu Lys	Pro Ala Lys Lys Thr Ser Asn Ser Ser	Asn Ser Ser Met	
35	40	45	
Gln Thr Pro	Thr Asn Ser Thr Thr Pro Glu Phe Leu Asn Gln Pro		
50	55	60	

(2) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1024 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 115...921
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

AGTTGGCAAA AACGCAGAGA CAGTAACGCA AAGGCAAATA AAGAGACTCA TTTTAAACAA	60
CGCAATGCCA TTACAAATAT AATCAGATCA GTTGGTGGGT TTTTACAAA GATT ATG	117
Met	
1	
AAG AGA GTT AGA GAA CTT GTA AAA AAA CAT CCC GAG AAA AGC AGT GTG	165
Lys Arg Val Arg Glu Leu Val Lys Lys His Pro Glu Lys Ser Ser Val	
5 10 15	
GCA TTA GTA GTA TTA ACC CAT GCT GCA TGC AAG AAA GCG AAA GAA TTG	213
Ala Leu Val Val Leu Thr His Ala Ala Cys Lys Lys Ala Lys Glu Leu	
20 25 30	
GAC GAT AAA GTC CAG GAT AAA TCC AAA CAA GCT GAA AAA GAA AAT CAA	261
Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu Asn Gln	
35 40 45	
ATC AAT TGG TGG AAA TAT TCA GGA TTA ACA ATA GCG ACA AGT TTA TTA	309
Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser Leu Leu	
50 55 60 65	
TTA GCC GCT TGT AGT GTT GGT GAT ATT GAT AAA CAG ATA GAG TTA GAA	357
Leu Ala Ala Cys Ser Val Gly Asp Ile Asp Lys Gln Ile Glu Leu Glu	
70 75 80	
CAA GAA AAA AAG GAA GCT GAA AAC GCT AGG GAT AGA GCG AAC AAG AGT	405
Gln Glu Lys Lys Glu Ala Glu Asn Ala Arg Asp Arg Ala Asn Lys Ser	
85 90 95	

-285-

GGG ATA GAA CTG GAA CAG GAA AAA CAA AAG ACC ATT AAA GAA CAA AAA	453
Gly Ile Glu Leu Glu Gln Glu Lys Gln Lys Thr Ile Lys Glu Gln Lys	
100 105 110	
GAT TTA GTT AAA AAA GCA GAA CAA AAT TGC CAA GAA AAT CAT GGC CAA	501
Asp Leu Val Lys Lys Ala Glu Gln Asn Cys Gln Glu Asn His Gly Gln	
115 120 125	
TTC TTT ATG AAA AAA TTA GGA ATT AAG GGT GGC ATT GCT ATA GAA GTA	549
Phe Phe Met Lys Lys Leu Gly Ile Lys Gly Gly Ile Ala Ile Glu Val	
130 135 140 145	
GAA GCT GAA TGC AAA ACC CCT AAA CCT GCA AAA ACC AAT CAA ACC CCT	597
Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr Asn Gln Thr Pro	
150 155 160	
ATC CAG CCA AAA CAC CTC CCC AAC TCT AAA CAA CCC CAC TCT CAA AGA	645
Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln Pro His Ser Gln Arg	
165 170 175	
GGA TCA AAA GCG CAA GAG CTT ATC GCT TAT TTG CAA AAA GAG TTA GAA	693
Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr Leu Gln Lys Glu Leu Glu	
180 185 190	
TCT CTG CCC TAT TCA CAA AAA GCT ATC GCT AAA CAA GTG AAT TTT TAC	741
Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys Gln Val Asn Phe Tyr	
195 200 205	
AGG CCA AGT TCT GTC GCT TAT TTA GAA CTA GAC CCT AGA GAT TTT AAG	789
Arg Pro Ser Ser Val Ala Tyr Leu Glu Leu Asp Pro Arg Asp Phe Lys	
210 215 220 225	
GTT ACA GAA GAA TGG CAA AAA GAA AAT CTA AAA ATA CGC TCT AAA GCT	837
Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys Ile Arg Ser Lys Ala	
230 235 240	
CAA GCT AAA ATG CTT GGA AAT GAG AAA CCC ACA AGC CCA CCT TTC AAC	885
Gln Ala Lys Met Leu Gly Asn Glu Lys Pro Thr Ser Pro Pro Phe Asn	
245 250 255	
CTC TCA AAG CCT TTT GTT CGT TCA AAA AAT ATT TGC TGATGTTAAT AAAGAA	937
Leu Ser Lys Pro Phe Val Arg Ser Lys Asn Ile Cys	
260 265	
ATAGAAGCAG TTGCTAATAC TGAAAAGAAA GCAGAAAAAG MGGGTTATGG TTATAGTAAA	997
AGGATGTAGG CATAAGAAAA TAAGAAC	1024

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 269 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-286-

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

```

Met Lys Arg Val Arg Glu Leu Val Lys Lys His Pro Glu Lys Ser Ser
 1           5           10           15
Val Ala Leu Val Val Leu Thr His Ala Ala Cys Lys Lys Ala Lys Glu
 20           25           30
Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu Asn
 35           40           45
Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Thr Ser Leu
 50           55           60
Leu Leu Ala Ala Cys Ser Val Gly Asp Ile Asp Lys Gln Ile Glu Leu
 65           70           75           80
Glu Gln Glu Lys Lys Glu Ala Glu Asn Ala Arg Asp Arg Ala Asn Lys
 85           90           95
Ser Gly Ile Glu Leu Glu Gln Glu Lys Gln Lys Thr Ile Lys Glu Gln
100           105           110
Lys Asp Leu Val Lys Lys Ala Glu Gln Asn Cys Gln Glu Asn His Gly
115           120           125
Gln Phe Phe Met Lys Lys Leu Gly Ile Lys Gly Gly Ile Ala Ile Glu
130           135           140
Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr Asn Gln Thr
145           150           155           160
Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln Pro His Ser Gln
165           170           175
Arg Gly Ser Lys Ala Gln Glu Leu Ile Ala Tyr Leu Gln Lys Glu Leu
180           185           190
Glu Ser Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys Gln Val Asn Phe
195           200           205
Tyr Arg Pro Ser Ser Val Ala Tyr Leu Glu Leu Asp Pro Arg Asp Phe
210           215           220
Lys Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys Ile Arg Ser Lys
225           230           235           240
Ala Gln Ala Lys Met Leu Gly Asn Glu Lys Pro Thr Ser Pro Pro Phe
245           250           255
Asn Leu Ser Lys Pro Phe Val Arg Ser Lys Asn Ile Cys
260           265

```

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 88...603
- (D) OTHER INFORMATION:

-287-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

```

AAAATAAGGA GGAATTGTTT GATTTTACGA TTGGCTGGAG CAAGCGTTTT AACGGCTTGT      60
GTCTTTTTCGG GGTGTTTTTT TTTAAAA ATG TTT GAT AAA AAA CTT TCT AGT AAC      114
      Met Phe Asp Lys Lys Leu Ser Ser Asn
      1              5

GAT TGG CAT ATC CAA AAA GTG GAA ATG AAC CAT CAA GTC TAT GAC ATT      162
Asp Trp His Ile Gln Lys Val Glu Met Asn His Gln Val Tyr Asp Ile
10              15              20              25

GAA ACC ATG CTC GCT GAT AGC GCT TTT AGA GAG CAT GAA GAA GAG CAA      210
Glu Thr Met Leu Ala Asp Ser Ala Phe Arg Glu His Glu Glu Glu Gln
      30              35              40

GAT TCC TCT CTA AAT ACC GCT TTG CCT GAA GAT AAA ACA GCG ATT GAA      258
Asp Ser Ser Leu Asn Thr Ala Leu Pro Glu Asp Lys Thr Ala Ile Glu
      45              50              55

GCC AAA GAG CAA GAG CAA AAA GAA AAA AGA AAA CGC TGG TAT GAG CTT      306
Ala Lys Glu Gln Glu Gln Lys Glu Lys Arg Lys Arg Trp Tyr Glu Leu
      60              65              70

TTT AAA AAG AAA CCA AAG CCC AAA AGC TCT ATG GGA GAG TTT GTG TTT      354
Phe Lys Lys Lys Pro Lys Pro Lys Ser Ser Met Gly Glu Phe Val Phe
      75              80              85

GAT CAA AAA GAA AAT CGT ATT TAT GGC AAA GGC TAT TGC AAC CGG TAT      402
Asp Gln Lys Glu Asn Arg Ile Tyr Gly Lys Gly Tyr Cys Asn Arg Tyr
90              95              100              105

TTT GCC AGC TAT GTA TGG CAG GGC GAT AGG CAC ATT GGG ATT GAA GAT      450
Phe Ala Ser Tyr Val Trp Gln Gly Asp Arg His Ile Gly Ile Glu Asp
      110              115              120

AGC GGG ATT TCA AGA AAA GTG TGT AAA GAT GAG CAT TTA ATG GCG TTT      498
Ser Gly Ile Ser Arg Lys Val Cys Lys Asp Glu His Leu Met Ala Phe
      125              130              135

GAA TTG GAA TTT ATG GAG AAT TTT AAG GGT AAT TTT ACG GTA ACT AAG      546
Glu Leu Glu Phe Met Glu Asn Phe Lys Gly Asn Phe Thr Val Thr Lys
      140              145              150

GGC AAG GAC ACG CTC ATT TTA GAC AAC CAA AAA ATG AAA ATT TAT TTG      594
Gly Lys Asp Thr Leu Ile Leu Asp Asn Gln Lys Met Lys Ile Tyr Leu
      155              160              165

AAA ACG CCT TGAGTGGGTT TTTGATTTC AACAATCTA AGATCACTAA ATTAGGGAT      652
Lys Thr Pro
170

TAAAAAGAAA TTTTTAA      669

```


-288-

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

```

Met Phe Asp Lys Lys Leu Ser Ser Asn Asp Trp His Ile Gln Lys Val
 1           5           10           15
Glu Met Asn His Gln Val Tyr Asp Ile Glu Thr Met Leu Ala Asp Ser
      20           25           30
Ala Phe Arg Glu His Glu Glu Glu Gln Asp Ser Ser Leu Asn Thr Ala
      35           40           45
Leu Pro Glu Asp Lys Thr Ala Ile Glu Ala Lys Glu Gln Glu Gln Lys
      50           55           60
Glu Lys Arg Lys Arg Trp Tyr Glu Leu Phe Lys Lys Lys Pro Lys Pro
      65           70           75           80
Lys Ser Ser Met Gly Glu Phe Val Phe Asp Gln Lys Glu Asn Arg Ile
      85           90           95
Tyr Gly Lys Gly Tyr Cys Asn Arg Tyr Phe Ala Ser Tyr Val Trp Gln
      100          105          110
Gly Asp Arg His Ile Gly Ile Glu Asp Ser Gly Ile Ser Arg Lys Val
      115          120          125
Cys Lys Asp Glu His Leu Met Ala Phe Glu Leu Glu Phe Met Glu Asn
      130          135          140
Phe Lys Gly Asn Phe Thr Val Thr Lys Gly Lys Asp Thr Leu Ile Leu
      145          150          155          160
Asp Asn Gln Lys Met Lys Ile Tyr Leu Lys Thr Pro
      165          170

```

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 87...1280
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

-289-

ATCAATCTAA CTTGAGTGGG TTTTTCGTAT TAGTTTCCAT GATATAATTT TGAAAAGTAA 60
 GATTGTTTTT TAAAAAAGG TTGGTA ATG GAA TCA GTA AAA ACA GGA AAA ACA 113
 Met Glu Ser Val Lys Thr Gly Lys Thr
 1 5

AAT AAG GTT GGC AAG AAT ACA GAG ATG GCT AAT ACA AAG GCA AAT AAA 161
 Asn Lys Val Gly Lys Asn Thr Glu Met Ala Asn Thr Lys Ala Asn Lys
 10 15 20 25

GAG ACT CAT TTT AAA CAA GTG AGC GCC ATT ACA AAT ATA ATC AGA TCA 209
 Glu Thr His Phe Lys Gln Val Ser Ala Ile Thr Asn Ile Ile Arg Ser
 30 35 40

GTT GGT GGG TTT TTT ACA AAA ATT GCA AAG AGA GTT AGA GGA CTT GTA 257
 Val Gly Gly Phe Phe Thr Lys Ile Ala Lys Arg Val Arg Gly Leu Val
 45 50 55

AAA AAA CAC CCC AAG AAA AGC AGT GCG GCA TTA GTA GTA TTG ACC CAT 305
 Lys Lys His Pro Lys Lys Ser Ser Ala Ala Leu Val Val Leu Thr His
 60 65 70

ATT GCG TGC AAG AAA GCG AAA GAA TTA GAC GAT AAA GTC CAA GAT AAA 353
 Ile Ala Cys Lys Lys Ala Lys Glu Leu Asp Asp Lys Val Gln Asp Lys
 75 80 85

TCC AAA CAA GCT GAA AAA GAA AAT CAA ATC AAT TGG TGG AAA TAT TCA 401
 Ser Lys Gln Ala Glu Lys Glu Asn Gln Ile Asn Trp Trp Lys Tyr Ser
 90 95 100 105

GGA TTA ACA ATA GCG GCA AGT TTA TTA TTA GCC GCT TGT AGC GCT GGT 449
 Gly Leu Thr Ile Ala Ala Ser Leu Leu Leu Ala Ala Cys Ser Ala Gly
 110 115 120

GAT ACT GAT AAA CAG ATA GAA CTA GAA CAA GAA AAA AAG GAA GCT GAA 497
 Asp Thr Asp Lys Gln Ile Glu Leu Glu Gln Glu Lys Lys Glu Ala Glu
 125 130 135

AAC GCT AGG GAT AGA GCG AAC AAG AGT GGG ATA GAA CTA GAA CAA GAA 545
 Asn Ala Arg Asp Arg Ala Asn Lys Ser Gly Ile Glu Leu Glu Gln Glu
 140 145 150

AGA CAG AAA ACA AAC AAG AGT GGG ATA GAA CTC GCT AAT AGT CAA ATA 593
 Arg Gln Lys Thr Asn Lys Ser Gly Ile Glu Leu Ala Asn Ser Gln Ile
 155 160 165

AAA GCA GAA CAA GAA AGA CAA AAG ACA GAA CAA GAA AAA CAA AAA GCA 641
 Lys Ala Glu Gln Glu Arg Gln Lys Thr Glu Gln Glu Lys Gln Lys Ala
 170 175 180 185

AAT AAG AGT GCG ATA GAG TTA GAA CAG CAA AAA CAA AAG ACC ATT AAT 689
 Asn Lys Ser Ala Ile Glu Leu Glu Gln Gln Lys Gln Lys Thr Ile Asn
 190 195 200

ACA CAA AGA GAT TTG ATT AAA GAA CAG AAA GAT TTC ATT AAA GAA ACA 737
 Thr Gln Arg Asp Leu Ile Lys Glu Gln Lys Asp Phe Ile Lys Glu Thr

-290-

205	210	215	
GAA CAA AAT TGC CAA GAA AAT CAT AAT CAA TTC TTT ATT AAA AAA TTA			785
Glu Gln Asn Cys Gln Glu Asn His Asn Gln Phe Phe Ile Lys Lys Leu			
220	225	230	
GGA ATT AAG GGT GGC ATT GCT ATA GAA GTA GAA GCT GAA TGC AAA ACC			833
Gly Ile Lys Gly Gly Ile Ala Ile Glu Val Glu Ala Glu Cys Lys Thr			
235	240	245	
CCT AAA CCT GCA AAA ACC AAT CAA ACC CCT ATC CAG CCA AAA CAC CTC			881
Pro Lys Pro Ala Lys Thr Asn Gln Thr Pro Ile Gln Pro Lys His Leu			
250	255	260	265
CCA AAC TCT AAA CAA CCT CAT TCT CAA AGA GGA TCA AAA GCG CAA GAG			929
Pro Asn Ser Lys Gln Pro His Ser Gln Arg Gly Ser Lys Ala Gln Glu			
270	275	280	
TTT ATC GCT TAT TTG CAA AAA GAG CTA GAA TTT CTG CCC TAT TCG CAA			977
Phe Ile Ala Tyr Leu Gln Lys Glu Leu Glu Phe Leu Pro Tyr Ser Gln			
285	290	295	
AAA GCT ATC GCT AAA CAA GTG AAT TTC TAT AAA CCA AGT TCT ATC GCT			1025
Lys Ala Ile Ala Lys Gln Val Asn Phe Tyr Lys Pro Ser Ser Ile Ala			
300	305	310	
TAT TTA GAA CTA GAT CCT AGA GAT TTT AAG GTT ACA GAA GAA TGG CAA			1073
Tyr Leu Glu Leu Asp Pro Arg Asp Phe Lys Val Thr Glu Glu Trp Gln			
315	320	325	
AAA GAA AAT CTA AAA ATA CGC TCT AAA GCT CAA GCT AAA ATG CTT GAA			1121
Lys Glu Asn Leu Lys Ile Arg Ser Lys Ala Gln Ala Lys Met Leu Glu			
330	335	340	345
ATG AGG GAT TTA AAA CCA GAC CCA CAA GCC CAC CTT CCA ACC TCT CAA			1169
Met Arg Asp Leu Lys Pro Asp Pro Gln Ala His Leu Pro Thr Ser Gln			
350	355	360	
AGC CTT TTG TTC GTT CAA AAA ATA TTT GCT GAT GTT AAT AAA GAA ATA			1217
Ser Leu Leu Phe Val Gln Lys Ile Phe Ala Asp Val Asn Lys Glu Ile			
365	370	375	
GAA GCA GTT GCT AAT ACT GAA AAG AAA GCA GAA AAA GCG GGT TAT GGT			1265
Glu Ala Val Ala Asn Thr Glu Lys Lys Ala Glu Lys Ala Gly Tyr Gly			
380	385	390	
TAT AGT AAA AGG ATG TAGGCATAAG AAAATAAGAA CACCATAAAA TCGTTTTTAG C			1321
Tyr Ser Lys Arg Met			
395			
TTCTAGGAGA CATCAGTCAG TTTCTTGCC			1350

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 398 amino acids

-291-

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

```

Met Glu Ser Val Lys Thr Gly Lys Thr Asn Lys Val Gly Lys Asn Thr
 1           5           10           15
Glu Met Ala Asn Thr Lys Ala Asn Lys Glu Thr His Phe Lys Gln Val
          20           25           30
Ser Ala Ile Thr Asn Ile Ile Arg Ser Val Gly Gly Phe Phe Thr Lys
          35           40           45
Ile Ala Lys Arg Val Arg Gly Leu Val Lys Lys His Pro Lys Lys Ser
          50           55           60
Ser Ala Ala Leu Val Val Leu Thr His Ile Ala Cys Lys Lys Ala Lys
          65           70           75           80
Glu Leu Asp Asp Lys Val Gln Asp Lys Ser Lys Gln Ala Glu Lys Glu
          85           90           95
Asn Gln Ile Asn Trp Trp Lys Tyr Ser Gly Leu Thr Ile Ala Ala Ser
          100          105          110
Leu Leu Leu Ala Ala Cys Ser Ala Gly Asp Thr Asp Lys Gln Ile Glu
          115          120          125
Leu Glu Gln Glu Lys Lys Glu Ala Glu Asn Ala Arg Asp Arg Ala Asn
          130          135          140
Lys Ser Gly Ile Glu Leu Glu Gln Glu Arg Gln Lys Thr Asn Lys Ser
          145          150          155          160
Gly Ile Glu Leu Ala Asn Ser Gln Ile Lys Ala Glu Gln Glu Arg Gln
          165          170          175
Lys Thr Glu Gln Glu Lys Gln Lys Ala Asn Lys Ser Ala Ile Glu Leu
          180          185          190
Glu Gln Gln Lys Gln Lys Thr Ile Asn Thr Gln Arg Asp Leu Ile Lys
          195          200          205
Glu Gln Lys Asp Phe Ile Lys Glu Thr Glu Gln Asn Cys Gln Glu Asn
          210          215          220
His Asn Gln Phe Phe Ile Lys Lys Leu Gly Ile Lys Gly Gly Ile Ala
          225          230          235          240
Ile Glu Val Glu Ala Glu Cys Lys Thr Pro Lys Pro Ala Lys Thr Asn
          245          250          255
Gln Thr Pro Ile Gln Pro Lys His Leu Pro Asn Ser Lys Gln Pro His
          260          265          270
Ser Gln Arg Gly Ser Lys Ala Gln Glu Phe Ile Ala Tyr Leu Gln Lys
          275          280          285
Glu Leu Glu Phe Leu Pro Tyr Ser Gln Lys Ala Ile Ala Lys Gln Val
          290          295          300
Asn Phe Tyr Lys Pro Ser Ser Ile Ala Tyr Leu Glu Leu Asp Pro Arg
          305          310          315          320
Asp Phe Lys Val Thr Glu Glu Trp Gln Lys Glu Asn Leu Lys Ile Arg
          325          330          335
Ser Lys Ala Gln Ala Lys Met Leu Glu Met Arg Asp Leu Lys Pro Asp
          340          345          350
Pro Gln Ala His Leu Pro Thr Ser Gln Ser Leu Leu Phe Val Gln Lys
          355          360          365

```

-292-

Ile Phe Ala Asp Val Asn Lys Glu Ile Glu Ala Val Ala Asn Thr Glu
 370 375 380
 Lys Lys Ala Glu Lys Ala Gly Tyr Gly Tyr Ser Lys Arg Met
 385 390 395

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 709 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 336...443
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

TAAGGGATAT TGCTAACGAT TAAGCTGTAT TGGAGAGATT TATTTTGCAA GAATTAATCT	60
TGCCTTGTGT GATTAGTAAC ACAAGGCAAG TGTGATAAAC CCTACTACAA TTTCAATTCA	120
AGGAGCCTAA CTAAAATAAA ATGAACAATT TCAGTTAGGG CTTTATTATA GCAAAAATTA	180
TCTAAGATTA CAAAGGGTAG CGTTTCTGTT TTTGGATTTA GAGCGTTATT TTGATTGTTT	240
TGAGTTTAAT TTACTTTTTG TTTAATAATA AATCTTAACT ATCATAAATG TACAATTAAA	300
GTATTTAAAA AAATTTTAAA ACAAAGGAT ATAAA ATG AAA ACC ATT AGA AAT	353
Met Lys Thr Ile Arg Asn	
1 5	
AGC GTG TTT ATT GGA GCG TCT TTA CTC GGC GGT TGC GCT AGC GTT GAG	401
Ser Val Phe Ile Gly Ala Ser Leu Leu Gly Gly Cys Ala Ser Val Glu	
10 15 20	
GCT TAT TTT GAC GCT TTG CAT GTT GCT CGC GTT AAA GAC GCT TGTTTATAG	452
Ala Tyr Phe Asp Ala Leu His Val Ala Arg Val Lys Asp Ala	
25 30 35	
AAAAAGAAGC ACACCACACG CCCAAAGACT TTGATAGCCC TTACCACACT GACTAAACCG	512
GCTAGAGTT TTAGTTGGGG GTTTTTAGGG GTGTTATTTT AGATACTCTC TGTTCCCTTA	572
AAGAAAATAA ATTTCTACCA TAAAATAAAA TCTTAAATTA AGGCGACTAA AACCCCACTT	632
TTAAAAAATT AAAAAGCGTT AAGTAAGACT TATCCAAAAA GCAAAGAAAA TCAATTTTTC	692
CAACCACTTT TTTTAAG	709

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-293-

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

```

Met Lys Thr Ile Arg Asn Ser Val Phe Ile Gly Ala Ser Leu Leu Gly
 1           5           10           15
Gly Cys Ala Ser Val Glu Ala Tyr Phe Asp Ala Leu His Val Ala Arg
          20           25           30
Val Lys Asp Ala
          35

```

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 888 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 19...837

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

```

AGATAGGAAT GTAAAGGA ATG GAA TTT ATG AAA AAG TTT GTA GCT TTA GGG      51
          Met Glu Phe Met Lys Lys Phe Val Ala Leu Gly
              1           5           10

CTT CTA TCC GCA GTT TTA AGC TCT TCG TTG TTA GCC GAA GGT GAT GGT      99
Leu Leu Ser Ala Val Leu Ser Ser Ser Leu Leu Ala Glu Gly Asp Gly
          15           20           25

GTT TAT ATA GGG ACT AAT TAT CAG CTT GGA CAA GCC CGT TTG AAT AGT     147
Val Tyr Ile Gly Thr Asn Tyr Gln Leu Gly Gln Ala Arg Leu Asn Ser
          30           35           40

AAT ATT TAT AAT ACA GGG GAT TGC ACA GGG AGT GTT GTA GGT TGC CCC     195
Asn Ile Tyr Asn Thr Gly Asp Cys Thr Gly Ser Val Val Gly Cys Pro
          45           50           55

CCA GGT CTT ACC GCT AAT AAG CAT AAT CCA GGA GGC ACC AAT ATC AAT     243
Pro Gly Leu Thr Ala Asn Lys His Asn Pro Gly Gly Thr Asn Ile Asn
          60           65           70           75

TGG CAT GCT AAA TAC GCT AAT GGG GCT TTG AAT GGT CTT GGG TTG AAT     291
Trp His Ala Lys Tyr Ala Asn Gly Ala Leu Asn Gly Leu Gly Leu Asn
          80           85           90

```

-294-

GTG GGT TAT AAG AAG TTC TTC CAG TTC AAG TCT TTT GAT ATG ACA AGC	339
Val Gly Tyr Lys Lys Phe Phe Gln Phe Lys Ser Phe Asp Met Thr Ser	
95 100 105	
AAG TGG TTT GGT TTT AGA GTG TAT GGG CTT TTT GAT TAT GGG CAT GCC	387
Lys Trp Phe Gly Phe Arg Val Tyr Gly Leu Phe Asp Tyr Gly His Ala	
110 115 120	
ACT TTA GGC AAG CAA GTT TAT GCA CCT AAT AAA ATC CAG TTG GAT ATG	435
Thr Leu Gly Lys Gln Val Tyr Ala Pro Asn Lys Ile Gln Leu Asp Met	
125 130 135	
GTC TCT TGG GGT GTG GGG AGC GAT TTG TTA GCT GAT ATT ATT GAT AAC	483
Val Ser Trp Gly Val Gly Ser Asp Leu Leu Ala Asp Ile Ile Asp Asn	
140 145 150 155	
GAT AAC GCT TCT TTT GGT ATT TTT GGT GGG GTC GCT ATC GGC GGT AAC	531
Asp Asn Ala Ser Phe Gly Ile Phe Gly Gly Val Ala Ile Gly Gly Asn	
160 165 170	
ACT TGG AAA AGC TCA GCG GCA AAC TAT TGG AAA GAG CAA ATC ATT GAA	579
Thr Trp Lys Ser Ser Ala Ala Asn Tyr Trp Lys Glu Gln Ile Ile Glu	
175 180 185	
GCT AAG GGT CCT GAT GTT TGT ACC CCT ACT TAT TGT AAC CCT AAC GCT	627
Ala Lys Gly Pro Asp Val Cys Thr Pro Thr Tyr Cys Asn Pro Asn Ala	
190 195 200	
CCT TAT AGC ACC AAA ACT TCA ACC GTC GCT TTT CAG GTA TGG TTG AAT	675
Pro Tyr Ser Thr Lys Thr Ser Thr Val Ala Phe Gln Val Trp Leu Asn	
205 210 215	
TTT GGG GTG AGA GCC AAT ATT TAC AAG CAT AAT GGC GTA GAG TTT GGC	723
Phe Gly Val Arg Ala Asn Ile Tyr Lys His Asn Gly Val Glu Phe Gly	
220 225 230 235	
GTG AGA GTG CCG CTA CTC ATC AAC AAG TTT TTG AGT GCG GGT CCT AAC	771
Val Arg Val Pro Leu Leu Ile Asn Lys Phe Leu Ser Ala Gly Pro Asn	
240 245 250	
GCT ACT AAT CTT TAT TAC CAT TTG AAA CGG GAT TAT TCG CTT TAT TTA	819
Ala Thr Asn Leu Tyr Tyr His Leu Lys Arg Asp Tyr Ser Leu Tyr Leu	
255 260 265	
GGG TAT AAC TAC ACT TTT TAAACCCTTT AAAAGGGTGT CTTTAAGCCC TTTTGTAGT	875
Gly Tyr Asn Tyr Thr Phe	
270	
CCTTATAAAA AGG	888

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 273 amino acids

(B) TYPE: amino acid

-295-

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

```

Met Glu Phe Met Lys Lys Phe Val Ala Leu Gly Leu Leu Ser Ala Val
 1           5           10           15
Leu Ser Ser Ser Leu Leu Ala Glu Gly Asp Gly Val Tyr Ile Gly Thr
      20           25           30
Asn Tyr Gln Leu Gly Gln Ala Arg Leu Asn Ser Asn Ile Tyr Asn Thr
      35           40           45
Gly Asp Cys Thr Gly Ser Val Val Gly Cys Pro Pro Gly Leu Thr Ala
      50           55           60
Asn Lys His Asn Pro Gly Gly Thr Asn Ile Asn Trp His Ala Lys Tyr
      65           70           75           80
Ala Asn Gly Ala Leu Asn Gly Leu Gly Leu Asn Val Gly Tyr Lys Lys
      85           90           95
Phe Phe Gln Phe Lys Ser Phe Asp Met Thr Ser Lys Trp Phe Gly Phe
      100          105          110
Arg Val Tyr Gly Leu Phe Asp Tyr Gly His Ala Thr Leu Gly Lys Gln
      115          120          125
Val Tyr Ala Pro Asn Lys Ile Gln Leu Asp Met Val Ser Trp Gly Val
      130          135          140
Gly Ser Asp Leu Leu Ala Asp Ile Ile Asp Asn Asp Asn Ala Ser Phe
      145          150          155          160
Gly Ile Phe Gly Gly Val Ala Ile Gly Gly Asn Thr Trp Lys Ser Ser
      165          170          175
Ala Ala Asn Tyr Trp Lys Glu Gln Ile Ile Glu Ala Lys Gly Pro Asp
      180          185          190
Val Cys Thr Pro Thr Tyr Cys Asn Pro Asn Ala Pro Tyr Ser Thr Lys
      195          200          205
Thr Ser Thr Val Ala Phe Gln Val Trp Leu Asn Phe Gly Val Arg Ala
      210          215          220
Asn Ile Tyr Lys His Asn Gly Val Glu Phe Gly Val Arg Val Pro Leu
      225          230          235          240
Leu Ile Asn Lys Phe Leu Ser Ala Gly Pro Asn Ala Thr Asn Leu Tyr
      245          250          255
Tyr His Leu Lys Arg Asp Tyr Ser Leu Tyr Leu Gly Tyr Asn Tyr Thr
      260          265          270
Phe

```

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 310 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

-296-

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 10...279
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

```

AAAAGGAGA GTG GCG GTG AAA AAA ATC GTT GTG AGT TGG TGT GTG GCG TTG      51
    Val Ala Val Lys Lys Ile Val Val Ser Trp Cys Val Ala Leu
        1             5             10

GCT TTT TTA AGC GCG GAT TCA GCA CAA GCC AAT AAA GCG ATC AGT AAT      99
Ala Phe Leu Ser Ala Asp Ser Ala Gln Ala Asn Lys Ala Ile Ser Asn
15             20             25             30

GCG GAT TTG ATT AAA GAG ATA AGG GAT TTA AAA AAA ATC ATC AGC GCG      147
Ala Asp Leu Ile Lys Glu Ile Arg Asp Leu Lys Lys Ile Ile Ser Ala
             35             40             45

CAA AAC ACT GAG ATT AAC AAC TTA AGA AAA GTG CAA GAA GTG TTG TCT      195
Gln Asn Thr Glu Ile Asn Asn Leu Arg Lys Val Gln Glu Val Leu Ser
             50             55             60

GGG CAA TTA GGG GAC ATG CGT AAG GAT ATA TTA AGC ACT AGA GAT TAT      243
Gly Gln Leu Gly Asp Met Arg Lys Asp Ile Leu Ser Thr Arg Asp Tyr
             65             70             75

TGC ATT AGC TTA AGG CCT TAT ATC TAT AAT TGG CGC TAGGGGATAA TCCAAA      295
Cys Ile Ser Leu Arg Pro Tyr Ile Tyr Asn Trp Arg
             80             85             90

AAATGAAAGC ATGCGC                                                    310

```

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

```

Val Ala Val Lys Lys Ile Val Val Ser Trp Cys Val Ala Leu Ala Phe
  1             5             10             15
Leu Ser Ala Asp Ser Ala Gln Ala Asn Lys Ala Ile Ser Asn Ala Asp
             20             25             30
Leu Ile Lys Glu Ile Arg Asp Leu Lys Lys Ile Ile Ser Ala Gln Asn
             35             40             45
Thr Glu Ile Asn Asn Leu Arg Lys Val Gln Glu Val Leu Ser Gly Gln
             50             55             60
Leu Gly Asp Met Arg Lys Asp Ile Leu Ser Thr Arg Asp Tyr Cys Ile
             65             70             75             80

```


-298-

His Asn Pro Ile Lys Asp Ala Arg Leu His Glu Cys Phe Glu Lys Ala
 125 130 135 140

CGC GCT TTG ATC CAC TAC CAA CGA AGT TTT TTT GAG GAA TGAATGGCTG AT 486
 Arg Ala Leu Ile His Tyr Gln Arg Ser Phe Phe Glu Glu
 145 150

TTATTGTCCA GTTTAAAAA CCTTCCTAAC AGCAGTGGCG TGTATCAATA TTTTGATAAA 546
 AAC 549

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

Met Glu Gln Asn Ile Phe Ser Leu Leu Ile Gln Lys Lys Ser Tyr Lys
 1 5 10 15
 Lys Leu Glu Thr Leu Leu Lys Leu Lys Lys Leu Lys Val Phe Met Pro
 20 25 30
 Leu Ser Leu Gln Glu Asn Leu Leu Phe Ile Phe Ile Lys Asp Ser Lys
 35 40 45
 Leu Leu Phe Ala Phe Lys Asp Ile Trp Ala Ser Lys Glu Phe Asn Gln
 50 55 60
 Arg Phe Ala Lys Glu Ile Ser His Phe Leu Asn Thr Gln Gly His Ala
 65 70 75 80
 Tyr Gly Phe Asp Gly Leu Asn Gly Leu Glu Ile Leu Gly Tyr Val Pro
 85 90 95
 Lys Asp Ala Leu Lys Lys Ser Asn Phe Tyr Ala Pro Ile Lys Lys Gln
 100 105 110
 Ala Arg Phe Phe Arg Pro Ser Ala Leu Gly Leu Phe His Asn Pro Ile
 115 120 125
 Lys Asp Ala Arg Leu His Glu Cys Phe Glu Lys Ala Arg Ala Leu Ile
 130 135 140
 His Tyr Gln Arg Ser Phe Phe Glu Glu
 145 150

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2627 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

-299-

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 18...2582

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

AAAGACATGT GCAACCG ATG AAA TCT AAA AAA CTT TAT TTG GCT TTA ATC	50
Met Lys Ser Lys Lys Leu Tyr Leu Ala Leu Ile	
1 5 10	
ATA GGG GTT TTA TTA GCG TTT TTA ACC CTA TCT TCA TGG CTG GGT AAT	98
Ile Gly Val Leu Leu Ala Phe Leu Thr Leu Ser Ser Trp Leu Gly Asn	
15 20 25	
AGC GGT TTA GTG GGG CGT TTT GGG GTG TGG TTT GCC GCA CTC AAT AAA	146
Ser Gly Leu Val Gly Arg Phe Gly Val Trp Phe Ala Ala Leu Asn Lys	
30 35 40	
AAA TAT TTT GGG CAT CTT TCA TTC ATT AAT TTA CCC TAT TTA GCA TGG	194
Lys Tyr Phe Gly His Leu Ser Phe Ile Asn Leu Pro Tyr Leu Ala Trp	
45 50 55	
GTT TTA TTC CTT TTA TAC AAG ACT AAA AAC CCT TTT ACA GAA ATC GTT	242
Val Leu Phe Leu Leu Tyr Lys Thr Lys Asn Pro Phe Thr Glu Ile Val	
60 65 70 75	
TTA GAA AAA ACT TTA GGG CAT CTA TTA GGC ATT TTA TCT TTG CTC TTT	290
Leu Glu Lys Thr Leu Gly His Leu Leu Gly Ile Leu Ser Leu Leu Phe	
80 85 90	
TTA CAA TCT AGC CTA TTA AAT CAA GGG GAA ATC GGC AAC AGC GCG CGT	338
Leu Gln Ser Ser Leu Leu Asn Gln Gly Glu Ile Gly Asn Ser Ala Arg	
95 100 105	
TTG TTT TTA CGC CCT TTT ATA GGG GAT TTT GGG CTT TAT GCG CTG ATA	386
Leu Phe Leu Arg Pro Phe Ile Gly Asp Phe Gly Leu Tyr Ala Leu Ile	
110 115 120	
ACG CTT ATG GTA GTT ATT TCT TAT TTG ATT CTA TTC AAA CTA CCC CCT	434
Thr Leu Met Val Val Ile Ser Tyr Leu Ile Leu Phe Lys Leu Pro Pro	
125 130 135	
AAA AGC GTT TTT TAT CCT TAT ATG AAC AAA ACA CAA AAC CTT TTA AAA	482
Lys Ser Val Phe Tyr Pro Tyr Met Asn Lys Thr Gln Asn Leu Leu Lys	
140 145 150 155	
GAG ATT TAC AAA CAA TGC TTA CAA GCC TTT AGC CCT AAT TTT AGC CCA	530
Glu Ile Tyr Lys Gln Cys Leu Gln Ala Phe Ser Pro Asn Phe Ser Pro	
160 165 170	
AAA AAA GAG GGT TTT GAA AAC ACC CCA TCA GAT ATT CAA AAA AAA GAA	578
Lys Lys Glu Gly Phe Glu Asn Thr Pro Ser Asp Ile Gln Lys Lys Glu	
175 180 185	

-300-

ACC AAA AAC GAC AAA GAA AAA GAA AAC CGC AAA GAA AAC CCT ATT AAT	626
Thr Lys Asn Asp Lys Glu Lys Glu Asn Arg Lys Glu Asn Pro Ile Asn	
190 195 200	
GAA AAC CAC AAA ACC CCT AAC GAA GAA CCG TTT TTA GCG ATC CCT ACC	674
Glu Asn His Lys Thr Pro Asn Glu Glu Pro Phe Leu Ala Ile Pro Thr	
205 210 215	
CCC TAT AAC ACG ACT TTA AAT GAT TCA GAG CCG CAA GAA GGC TTA GTC	722
Pro Tyr Asn Thr Thr Leu Asn Asp Ser Glu Pro Gln Glu Gly Leu Val	
220 225 230 235	
CAA ATT TCC TCC CAC CCC CCT ACC CAT TAC ACC ATT TAC CCT AAA AGA	770
Gln Ile Ser Ser His Pro Pro Thr His Tyr Thr Ile Tyr Pro Lys Arg	
240 245 250	
AAC CGA TTT GAT GAT TTG ACT AAC CCC ACT AAC CCC CCT TTA AAA GAA	818
Asn Arg Phe Asp Asp Leu Thr Asn Pro Thr Asn Pro Pro Leu Lys Glu	
255 260 265	
ATT AAA CAA GAA ACT AAA GAA AGA GAA CCC ACG CCT ACA AAA GAA ACT	866
Ile Lys Gln Glu Thr Lys Glu Arg Glu Pro Thr Pro Thr Lys Glu Thr	
270 275 280	
CTT ACG CCC ACC ACG CCC AAA CCT ATC ATG CCC ACA CTT GCA CCC ATA	914
Leu Thr Pro Thr Thr Pro Lys Pro Ile Met Pro Thr Leu Ala Pro Ile	
285 290 295	
ATA GAA AAT GAC AAC AAA ACA GAA AAC CAA AAA ACC CCC AAC CAC CCT	962
Ile Glu Asn Asp Asn Lys Thr Glu Asn Gln Lys Thr Pro Asn His Pro	
300 305 310 315	
AAA AAA GAA GAA AAC CCA CAA GAA AAC ACG CAA GAA GAA ATG ATA GAA	1010
Lys Lys Glu Glu Asn Pro Gln Glu Asn Thr Gln Glu Glu Met Ile Glu	
320 325 330	
GGA AGG ATA GAA GAA ATG ATA AAG GAA AAT CTA AAA AAA GAA GAA AAA	1058
Gly Arg Ile Glu Glu Met Ile Lys Glu Asn Leu Lys Lys Glu Glu Lys	
335 340 345	
GAA GTG CAA AAC GCT CCA AAC TTT AGC CCA GTA ACC CCC ACA AGC GCT	1106
Glu Val Gln Asn Ala Pro Asn Phe Ser Pro Val Thr Pro Thr Ser Ala	
350 355 360	
AAA AAA CCC GTT ATG GTT AAA GAA TTG AGC GAA AAT AAA GAG ATA TTA	1154
Lys Lys Pro Val Met Val Lys Glu Leu Ser Glu Asn Lys Glu Ile Leu	
365 370 375	
GAC GGA TTG GAT TAT GGC GAA GTG CAA AAA CCC AAA GAT TAT GAG CTT	1202
Asp Gly Leu Asp Tyr Gly Glu Val Gln Lys Pro Lys Asp Tyr Glu Leu	
380 385 390 395	
CCC ACC ACG CAA TTA TTG AAT GCG GTT TGT TTG AAA GAC ACT TCT TTA	1250
Pro Thr Thr Gln Leu Leu Asn Ala Val Cys Leu Lys Asp Thr Ser Leu	
400 405 410	

-301-

GAC GAA AAC GAG ATT GAC CAA AAA ATC CAG GAT CTA TTG AGC AAA CTG	1298
Asp Glu Asn Glu Ile Asp Gln Lys Ile Gln Asp Leu Leu Ser Lys Leu	
415 420 425	
CGC ACC TTT AAA ATT GAT GGC GAT ATT ATC CGC ACT TAT TCA GGC CCT	1346
Arg Thr Phe Lys Ile Asp Gly Asp Ile Ile Arg Thr Tyr Ser Gly Pro	
430 435 440	
ATT GTA ACC ACT TTT GAA TTC CGC CCA GCC CCT AAC GTT AAG GTG AGT	1394
Ile Val Thr Thr Phe Glu Phe Arg Pro Ala Pro Asn Val Lys Val Ser	
445 450 455	
CGT ATT TTA GGC TTG AGC GAT GAT TTA GCG ATG ACT TTA TGC GCT GAA	1442
Arg Ile Leu Gly Leu Ser Asp Asp Leu Ala Met Thr Leu Cys Ala Glu	
460 465 470 475	
TCC ATC CGC ATT CAA GCC CCT ATT AAG GGT AAA GAT GTC GTT GGC ATT	1490
Ser Ile Arg Ile Gln Ala Pro Ile Lys Gly Lys Asp Val Val Gly Ile	
480 485 490	
GAA ATC CCT AAC AGC CAA AGC CAA ATT ATT TAT TTA AGA GAA ATT CTA	1538
Glu Ile Pro Asn Ser Gln Ser Gln Ile Ile Tyr Leu Arg Glu Ile Leu	
495 500 505	
GAG AGC GAA TTG TTT CAA AAA TCC AGC TCG CCC TTA ACT CTA GCT TTA	1586
Glu Ser Glu Leu Phe Gln Lys Ser Ser Ser Pro Leu Thr Leu Ala Leu	
510 515 520	
GGC AAA GAC ATT GTG GGT AAC CCT TTC ATC ACG GAT TTA AAA AAG CTC	1634
Gly Lys Asp Ile Val Gly Asn Pro Phe Ile Thr Asp Leu Lys Lys Leu	
525 530 535	
CCC CAT TTG CTC ATC GCT GGC ACG ACA GGA AGC GGT AAG AGC GTG GGC	1682
Pro His Leu Leu Ile Ala Gly Thr Thr Gly Ser Gly Lys Ser Val Gly	
540 545 550 555	
GTG AAT GCG ATG ATT TTA TCC TTA CTT TAT AAA AAC CCT CCC GAT CAA	1730
Val Asn Ala Met Ile Leu Ser Leu Leu Tyr Lys Asn Pro Pro Asp Gln	
560 565 570	
CTC AAA TTA GTG ATG ATC GAT CCC AAA ATG GTA GAA TTT AGT ATT TAT	1778
Leu Lys Leu Val Met Ile Asp Pro Lys Met Val Glu Phe Ser Ile Tyr	
575 580 585	
GCG GAT ATC CCT CAT TTG CTC ACG CCC ATT ATC ACC GAC CCT AAA AAA	1826
Ala Asp Ile Pro His Leu Leu Thr Pro Ile Ile Thr Asp Pro Lys Lys	
590 595 600	
GCT ATT GGG GCT TTG CAA AGC GTG GCT AAA GAA ATG GAA CGC CGG TAT	1874
Ala Ile Gly Ala Leu Gln Ser Val Ala Lys Glu Met Glu Arg Arg Tyr	
605 610 615	
TCT TTA ATG AGC GAA TAC AAG GTT AAA ACC ATT GAT TCT TAT AAT GAA	1922
Ser Leu Met Ser Glu Tyr Lys Val Lys Thr Ile Asp Ser Tyr Asn Glu	
620 625 630 635	

-302-

CAA GCC CCA AGT AAC GGC GTT GAA GCG TTC CCC TAT TTG ATT GTG GTG	1970
Gln Ala Pro Ser Asn Gly Val Glu Ala Phe Pro Tyr Leu Ile Val Val	
640 645 650	
ATT GAT GAA TTA GCG GAT TTA ATG ATG ACA GGG GGC AAA GAA GCG GAG	2018
Ile Asp Glu Leu Ala Asp Leu Met Met Thr Gly Gly Lys Glu Ala Glu	
655 660 665	
TTT CCT ATC GCT AGA ATC GCT CAA ATG GGG CGC GCG AGC GGC TTA CAC	2066
Phe Pro Ile Ala Arg Ile Ala Gln Met Gly Arg Ala Ser Gly Leu His	
670 675 680	
CTC ATT GTA GCG ACC CAA CGC CCA AGC GTG GAT GTC GTA ACC GGC TTG	2114
Leu Ile Val Ala Thr Gln Arg Pro Ser Val Asp Val Val Thr Gly Leu	
685 690 695	
ATT AAA ACC AAC TTG CCT TCA AGG GTG AGT TTT AGG GTA GGC ACT AAG	2162
Ile Lys Thr Asn Leu Pro Ser Arg Val Ser Phe Arg Val Gly Thr Lys	
700 705 710 715	
ATT GAT TCT AAA GTG ATT TTA GAC ACT GAT GGG GCG CAA AGC TTG TTA	2210
Ile Asp Ser Lys Val Ile Leu Asp Thr Asp Gly Ala Gln Ser Leu Leu	
720 725 730	
GGA AGA GGC GAT ATG CTC TTT ACC CCC CCA GGA GCG AAC GGG TTA GTG	2258
Gly Arg Gly Asp Met Leu Phe Thr Pro Pro Gly Ala Asn Gly Leu Val	
735 740 745	
CGC TTG CAT GCC CCC TTT GCC ACT GAA GAT GAA ATC AAA AAA ATC GTG	2306
Arg Leu His Ala Pro Phe Ala Thr Glu Asp Glu Ile Lys Lys Ile Val	
750 755 760	
GAT TTT ATT AAA GCC CAA AAA GAA GTA CAA TAC GAT AAA GAT TTC TTG	2354
Asp Phe Ile Lys Ala Gln Lys Glu Val Gln Tyr Asp Lys Asp Phe Leu	
765 770 775	
CTA GAA GAA TCA CGC ATG CCT TTA GAC ACC CCT AAT TAT CAA GGC GAT	2402
Leu Glu Glu Ser Arg Met Pro Leu Asp Thr Pro Asn Tyr Gln Gly Asp	
780 785 790 795	
GAC ATT TTA GAA AGG GCT AAA GCG GTG ATT TTA GAA AAA AAG ATC ACT	2450
Asp Ile Leu Glu Arg Ala Lys Ala Val Ile Leu Glu Lys Lys Ile Thr	
800 805 810	
TCT ACG AGT TTT TTA CAA CGC CAA TTA AAA ATC GGC TAC AAC CAA GCC	2498
Ser Thr Ser Phe Leu Gln Arg Gln Leu Lys Ile Gly Tyr Asn Gln Ala	
815 820 825	
GCT ACC ATT ACT GAC GAA TTA GAA GCT CAA GGC TTT TTA TCC CCA AGA	2546
Ala Thr Ile Thr Asp Glu Leu Glu Ala Gln Gly Phe Leu Ser Pro Arg	
830 835 840	
AAC GCT AAA GGC AAC AGA GAG ATT TTG CAA AAC TTT TAGGCTTTGT TTTTAT	2598
Asn Ala Lys Gly Asn Arg Glu Ile Leu Gln Asn Phe	
845 850 855	

-303-

TGGATATTGG CAAACATAT TTTTGATT

2627

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 855 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

```

Met Lys Ser Lys Lys Leu Tyr Leu Ala Leu Ile Ile Gly Val Leu Leu
 1             5             10             15
Ala Phe Leu Thr Leu Ser Ser Trp Leu Gly Asn Ser Gly Leu Val Gly
 20             25             30
Arg Phe Gly Val Trp Phe Ala Ala Leu Asn Lys Lys Tyr Phe Gly His
 35             40             45
Leu Ser Phe Ile Asn Leu Pro Tyr Leu Ala Trp Val Leu Phe Leu Leu
 50             55             60
Tyr Lys Thr Lys Asn Pro Phe Thr Glu Ile Val Leu Glu Lys Thr Leu
 65             70             75             80
Gly His Leu Leu Gly Ile Leu Ser Leu Leu Phe Leu Gln Ser Ser Leu
 85             90             95
Leu Asn Gln Gly Glu Ile Gly Asn Ser Ala Arg Leu Phe Leu Arg Pro
100             105             110
Phe Ile Gly Asp Phe Gly Leu Tyr Ala Leu Ile Thr Leu Met Val Val
115             120             125
Ile Ser Tyr Leu Ile Leu Phe Lys Leu Pro Pro Lys Ser Val Phe Tyr
130             135             140
Pro Tyr Met Asn Lys Thr Gln Asn Leu Leu Lys Glu Ile Tyr Lys Gln
145             150             155             160
Cys Leu Gln Ala Phe Ser Pro Asn Phe Ser Pro Lys Lys Glu Gly Phe
165             170             175
Glu Asn Thr Pro Ser Asp Ile Gln Lys Lys Glu Thr Lys Asn Asp Lys
180             185             190
Glu Lys Glu Asn Arg Lys Glu Asn Pro Ile Asn Glu Asn His Lys Thr
195             200             205
Pro Asn Glu Glu Pro Phe Leu Ala Ile Pro Thr Pro Tyr Asn Thr Thr
210             215             220
Leu Asn Asp Ser Glu Pro Gln Glu Gly Leu Val Gln Ile Ser Ser His
225             230             235             240
Pro Pro Thr His Tyr Thr Ile Tyr Pro Lys Arg Asn Arg Phe Asp Asp
245             250             255
Leu Thr Asn Pro Thr Asn Pro Pro Leu Lys Glu Ile Lys Gln Glu Thr
260             265             270
Lys Glu Arg Glu Pro Thr Pro Thr Lys Glu Thr Leu Thr Pro Thr Thr
275             280             285
Pro Lys Pro Ile Met Pro Thr Leu Ala Pro Ile Ile Glu Asn Asp Asn
290             295             300
Lys Thr Glu Asn Gln Lys Thr Pro Asn His Pro Lys Lys Glu Glu Asn
305             310             315             320
Pro Gln Glu Asn Thr Gln Glu Glu Met Ile Glu Gly Arg Ile Glu Glu
325             330             335

```


BNSDOCID: <WO__9821225A1_I_>

-305-

770		775		780
Met Pro Leu Asp Thr	Pro Asn Tyr Gln Gly Asp	Asp Ile Leu Glu Arg		
785	790	795		800
Ala Lys Ala Val Ile	Leu Glu Lys Lys Ile Thr Ser Thr Ser Phe Leu			
	805	810		815
Gln Arg Gln Leu Lys Ile Gly Tyr Asn Gln Ala Ala Thr Ile Thr Asp				
	820	825		830
Glu Leu Glu Ala Gln Gly Phe Leu Ser Pro Arg Asn Ala Lys Gly Asn				
	835	840		845
Arg Glu Ile Leu Gln Asn Phe				
850	855			

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1986 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 56...1945
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

GGTGTCTCTTA AACAGCAGGG TGAAAGAGAT TTAAAAGAA AGCGCTCTGC ATTCT ATG	58
	Met
	1
CAA GAT AGT TTG CAT TTT AAG GTT AAT GAA GTG CAA GGG GTT TTA GAA	106
Gln Asp Ser Leu His Phe Lys Val Asn Glu Val Gln Gly Val Leu Glu	
5 10 15	
AAC ACT TAT ACG AGC ATG GGC ATT GTT AAA GAA ATG CTC CCT AAA GAC	154
Asn Thr Tyr Thr Ser Met Gly Ile Val Lys Glu Met Leu Pro Lys Asp	
20 25 30	
ACC AAA AGA GAA ATC AAA ATC GGC TTG TTA AAA AAC TTC ATT TTA GCC	202
Thr Lys Arg Glu Ile Lys Ile Gly Leu Leu Lys Asn Phe Ile Leu Ala	
35 40 45	
AAT TCG CAT GTC GCT GGG GTG AGC ATG TTT TTT AAA GGC AGA GAA GAT	250
Asn Ser His Val Ala Gly Val Ser Met Phe Phe Lys Gly Arg Glu Asp	
50 55 60 65	
TTA AGA TTA ACG CTT TTA AGG GAT AAC AAT ACG ATT AAG CTA GTG GAA	298
Leu Arg Leu Thr Leu Leu Arg Asp Asn Asn Thr Ile Lys Leu Val Glu	
70 75 80	

-306-

AAT CCG TCA TTA GAG AAT AGC CCT TTA GCG CAA AAA GCG ATG AAA AAT	346
Asn Pro Ser Leu Glu Asn Ser Pro Leu Ala Gln Lys Ala Met Lys Asn	
85 90 95	
AAA GAA ATT TCT AAA AGT TTG GGT TAT TAT AGG AAA ATG CCT AAT GGG	394
Lys Glu Ile Ser Lys Ser Leu Gly Tyr Tyr Arg Lys Met Pro Asn Gly	
100 105 110	
GCG GAA GTT TAT GGG GTG GAT ATT CTT TTA CCT TTA TTG AAT GAG AAC	442
Ala Glu Val Tyr Gly Val Asp Ile Leu Leu Pro Leu Leu Asn Glu Asn	
115 120 125	
GCT CAA GAG GTT GTA GGG GCT TTG ATG ATT TTT ATT TCC ATT GAC AGC	490
Ala Gln Glu Val Val Gly Ala Leu Met Ile Phe Ile Ser Ile Asp Ser	
130 135 140 145	
TTC AGC AAT GAA ATC ACT AAA AAC AGG AGC GAT TTA TTT TTA ATT GGC	538
Phe Ser Asn Glu Ile Thr Lys Asn Arg Ser Asp Leu Phe Leu Ile Gly	
150 155 160	
ACT AAA GGT AAA GTG CTT TTG AGC GCG AAT AAG AGT TTG CAA GAC AAA	586
Thr Lys Gly Lys Val Leu Leu Ser Ala Asn Lys Ser Leu Gln Asp Lys	
165 170 175	
CCT ATC GCA GAA ATT TAT AAG AGC GTG CCT AAA GCC ACC AAC GAA GTG	634
Pro Ile Ala Glu Ile Tyr Lys Ser Val Pro Lys Ala Thr Asn Glu Val	
180 185 190	
ATG GCT ATT TTA GAA AAC GGC TCT AAA GCG ACT TTA GAA TAC TTA GAT	682
Met Ala Ile Leu Glu Asn Gly Ser Lys Ala Thr Leu Glu Tyr Leu Asp	
195 200 205	
CCC TTT AGC CAT AAG GAA AAT TTT TTA GCC GTT GAA ACC TTT AAA ATG	730
Pro Phe Ser His Lys Glu Asn Phe Leu Ala Val Glu Thr Phe Lys Met	
210 215 220 225	
CTA GGC AAA ACA GAA AGT AAA GAC AAT CTT AAT TGG ATG ATC GCT TTA	778
Leu Gly Lys Thr Glu Ser Lys Asp Asn Leu Asn Trp Met Ile Ala Leu	
230 235 240	
ATC ATT GAA AAA GAC AAG GTC TAT GAG CAA GTA GGC TCG GTG CGT TTT	826
Ile Ile Glu Lys Asp Lys Val Tyr Glu Gln Val Gly Ser Val Arg Phe	
245 250 255	
GTG GTG ATC ATA GCG AGC GCA ATC ATG GTG TTA GCC TTG ATT ATA GCG	874
Val Val Ile Ile Ala Ser Ala Ile Met Val Leu Ala Leu Ile Ile Ala	
260 265 270	
ATC ACT CTC TTA ATG CGA GCG ATC GTG AGC AGT CGT TTG GAA GCC GTT	922
Ile Thr Leu Leu Met Arg Ala Ile Val Ser Ser Arg Leu Glu Ala Val	
275 280 285	
TCT AGC ACC TTG TCT CAT TTC TTT AAA TTA TTG AAC AAT CAA GCC AAT	970
Ser Ser Thr Leu Ser His Phe Phe Lys Leu Leu Asn Asn Gln Ala Asn	
290 295 300 305	

-307-

TCT AGC GGT ATT AAA TTG ATT GAA GCG AAA TCC AAT GAC GAG TTA GGC	1018
Ser Ser Gly Ile Lys Leu Ile Glu Ala Lys Ser Asn Asp Glu Leu Gly	
310 315 320	
CGC ATG CAA ACA GCG ATC AAT AAA AAT ATC TTG CAA ACC CAA AAA ATC	1066
Arg Met Gln Thr Ala Ile Asn Lys Asn Ile Leu Gln Thr Gln Lys Ile	
325 330 335	
ATG CAA GAA GAC AGG CAA GCC GTC CAA GAC ACC ATT AAA GTG GTT TCA	1114
Met Gln Glu Asp Arg Gln Ala Val Gln Asp Thr Ile Lys Val Val Ser	
340 345 350	
GAT GTG AAA GCA GGG AAT TTT GCG GTG CGC ATC ACG GCT GAG CCC GCA	1162
Asp Val Lys Ala Gly Asn Phe Ala Val Arg Ile Thr Ala Glu Pro Ala	
355 360 365	
AGC CCT GAT TTG AAA GAA TTG AGG GAC GCG CTA AAT GGG ATC ATG GAT	1210
Ser Pro Asp Leu Lys Glu Leu Arg Asp Ala Leu Asn Gly Ile Met Asp	
370 375 380 385	
TAT TTG CAA GAA AGC GTA GGG ACT CAC ATG CCA AGC ATT TTC AAA ATC	1258
Tyr Leu Gln Glu Ser Val Gly Thr His Met Pro Ser Ile Phe Lys Ile	
390 395 400	
TTT GAA AGC TAT TCT GGT TTG GAT TTT AGA GGC CGG ATC CAA AAC GCT	1306
Phe Glu Ser Tyr Ser Gly Leu Asp Phe Arg Gly Arg Ile Gln Asn Ala	
405 410 415	
TCG GGT AGG GTG GAA CTG GTT ACT AAC GCT TTA GGG CAA GAA ATC CAA	1354
Ser Gly Arg Val Glu Leu Val Thr Asn Ala Leu Gly Gln Glu Ile Gln	
420 425 430	
AAA ATG CTA GAA ACT TCG TCT AAT TTT GCC AAA GAT TTA GCG AAC GAT	1402
Lys Met Leu Glu Thr Ser Ser Asn Phe Ala Lys Asp Leu Ala Asn Asp	
435 440 445	
AGC GCG AAT TTA AAA GAG TGC GTG CAA AAT TTA GAA AAA GCT TCA AAC	1450
Ser Ala Asn Leu Lys Glu Cys Val Gln Asn Leu Glu Lys Ala Ser Asn	
450 455 460 465	
TCC CAA CAC AAA AGC TTG ATG GAA ACT TCC AAA ACG ATA GAA AAT ATC	1498
Ser Gln His Lys Ser Leu Met Glu Thr Ser Lys Thr Ile Glu Asn Ile	
470 475 480	
ACC ACT TCC ATT CAA GGC GTG AGC TCT CAA AGT GAA GCC ATG ATT GAA	1546
Thr Thr Ser Ile Gln Gly Val Ser Ser Gln Ser Glu Ala Met Ile Glu	
485 490 495	
CAA GGG CAA GAC ATT AAA AGC ATT GTA GAA ATC ATT AGA GAT ATT GCT	1594
Gln Gly Gln Asp Ile Lys Ser Ile Val Glu Ile Ile Arg Asp Ile Ala	
500 505 510	
GAT CAA ACC AAT CTT TTA GCC TTA AAC GCC GCT ATT GAA GCC GCA AGG	1642
Asp Gln Thr Asn Leu Leu Ala Leu Asn Ala Ala Ile Glu Ala Ala Arg	
515 520 525	

-308-

```

GCC GGC GAG CAT GGC AGA GGC TTT GCG GTG GTG GCT GAT GAG CTA AGA      1690
Ala Gly Glu His Gly Arg Gly Phe Ala Val Val Ala Asp Glu Val Arg
530                               535                               540                               545

AAG CTC GCT GAA AGG ACG CAA AAA TCG CTC AGC GAG ATT GAA GCC AAT      1738
Lys Leu Ala Glu Arg Thr Gln Lys Ser Leu Ser Glu Ile Glu Ala Asn
                               550                               555                               560

ATC AAT ATT TTA GTG CAA AGC ATT TCA GAC ACG AGC GAA AGC ATT AAA      1786
Ile Asn Ile Leu Val Gln Ser Ile Ser Asp Thr Ser Glu Ser Ile Lys
                               565                               570                               575

AAC CAG GTT AAA GAA GTG GAA GAA ATC AAC GCT TCT ATT GAA GCC TTA      1834
Asn Gln Val Lys Glu Val Glu Glu Ile Asn Ala Ser Ile Glu Ala Leu
                               580                               585                               590

AGA TCG GTT ACT GAG GGC AAT CTA AAA ATC GCT AGC GAT TCT TTA GAA      1882
Arg Ser Val Thr Glu Gly Asn Leu Lys Ile Ala Ser Asp Ser Leu Glu
                               595                               600                               605

ATC AGT CAA GAA ATT GAC AAA GTT TCT AAC GAT ATT TTA GAA GAT GTG      1930
Ile Ser Gln Glu Ile Asp Lys Val Ser Asn Asp Ile Leu Glu Asp Val
610                               615                               620                               625

AAT AAA AAG CAG TTT TAATGCTCAT TCATATTTGC TGCTCAGTGG ATAACCTCTA T  1986
Asn Lys Lys Gln Phe
                               630

```

1986

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 630 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

```

Met Gln Asp Ser Leu His Phe Lys Val Asn Glu Val Gln Gly Val Leu
 1             5             10             15
Glu Asn Thr Tyr Thr Ser Met Gly Ile Val Lys Glu Met Leu Pro Lys
20             25             30
Asp Thr Lys Arg Glu Ile Lys Ile Gly Leu Leu Lys Asn Phe Ile Leu
35             40             45
Ala Asn Ser His Val Ala Gly Val Ser Met Phe Phe Lys Gly Arg Glu
50             55             60
Asp Leu Arg Leu Thr Leu Leu Arg Asp Asn Asn Thr Ile Lys Leu Val
65             70             75             80
Glu Asn Pro Ser Leu Glu Asn Ser Pro Leu Ala Gln Lys Ala Met Lys
85             90             95
Asn Lys Glu Ile Ser Lys Ser Leu Gly Tyr Tyr Arg Lys Met Pro Asn

```

- 309 -

100					105					110					
Gly	Ala	Glu	Val	Tyr	Gly	Val	Asp	Ile	Leu	Leu	Pro	Leu	Leu	Asn	Glu
115					120					125					
Asn	Ala	Gln	Glu	Val	Val	Gly	Ala	Leu	Met	Ile	Phe	Ile	Ser	Ile	Asp
130					135					140					
Ser	Phe	Ser	Asn	Glu	Ile	Thr	Lys	Asn	Arg	Ser	Asp	Leu	Phe	Leu	Ile
145	150					155					160				
Gly	Thr	Lys	Gly	Lys	Val	Leu	Leu	Ser	Ala	Asn	Lys	Ser	Leu	Gln	Asp
165					170					175					
Lys	Pro	Ile	Ala	Glu	Ile	Tyr	Lys	Ser	Val	Pro	Lys	Ala	Thr	Asn	Glu
180					185					190					
Val	Met	Ala	Ile	Leu	Glu	Asn	Gly	Ser	Lys	Ala	Thr	Leu	Glu	Tyr	Leu
195					200					205					
Asp	Pro	Phe	Ser	His	Lys	Glu	Asn	Phe	Leu	Ala	Val	Glu	Thr	Phe	Lys
210					215					220					
Met	Leu	Gly	Lys	Thr	Glu	Ser	Lys	Asp	Asn	Leu	Asn	Trp	Met	Ile	Ala
225	230					235					240				
Leu	Ile	Ile	Glu	Lys	Asp	Lys	Val	Tyr	Glu	Gln	Val	Gly	Ser	Val	Arg
245					250					255					
Phe	Val	Val	Ile	Ile	Ala	Ser	Ala	Ile	Met	Val	Leu	Ala	Leu	Ile	Ile
260					265					270					
Ala	Ile	Thr	Leu	Leu	Met	Arg	Ala	Ile	Val	Ser	Ser	Arg	Leu	Glu	Ala
275					280					285					
Val	Ser	Ser	Thr	Leu	Ser	His	Phe	Phe	Lys	Leu	Leu	Asn	Asn	Gln	Ala
290					295					300					
Asn	Ser	Ser	Gly	Ile	Lys	Leu	Ile	Glu	Ala	Lys	Ser	Asn	Asp	Glu	Leu
305	310					315					320				
Gly	Arg	Met	Gln	Thr	Ala	Ile	Asn	Lys	Asn	Ile	Leu	Gln	Thr	Gln	Lys
325					330					335					
Ile	Met	Gln	Glu	Asp	Arg	Gln	Ala	Val	Gln	Asp	Thr	Ile	Lys	Val	Val
340					345					350					
Ser	Asp	Val	Lys	Ala	Gly	Asn	Phe	Ala	Val	Arg	Ile	Thr	Ala	Glu	Pro
355					360					365					
Ala	Ser	Pro	Asp	Leu	Lys	Glu	Leu	Arg	Asp	Ala	Leu	Asn	Gly	Ile	Met
370					375					380					
Asp	Tyr	Leu	Gln	Glu	Ser	Val	Gly	Thr	His	Met	Pro	Ser	Ile	Phe	Lys
385	390					395					400				
Ile	Phe	Glu	Ser	Tyr	Ser	Gly	Leu	Asp	Phe	Arg	Gly	Arg	Ile	Gln	Asn
405					410					415					
Ala	Ser	Gly	Arg	Val	Glu	Leu	Val	Thr	Asn	Ala	Leu	Gly	Gln	Glu	Ile
420					425					430					
Gln	Lys	Met	Leu	Glu	Thr	Ser	Ser	Asn	Phe	Ala	Lys	Asp	Leu	Ala	Asn
435					440					445					
Asp	Ser	Ala	Asn	Leu	Lys	Glu	Cys	Val	Gln	Asn	Leu	Glu	Lys	Ala	Ser
450					455					460					
Asn	Ser	Gln	His	Lys	Ser	Leu	Met	Glu	Thr	Ser	Lys	Thr	Ile	Glu	Asn
465	470					475					480				
Ile	Thr	Thr	Ser	Ile	Gln	Gly	Val	Ser	Ser	Gln	Ser	Glu	Ala	Met	Ile
485					490					495					
Glu	Gln	Gly	Gln	Asp	Ile	Lys	Ser	Ile	Val	Glu	Ile	Ile	Arg	Asp	Ile
500					505					510					
Ala	Asp	Gln	Thr	Asn	Leu	Leu	Ala	Leu	Asn	Ala	Ala	Ile	Glu	Ala	Ala
515					520					525					
Arg	Ala	Gly	Glu	His	Gly	Arg	Gly	Phe	Ala	Val	Val	Ala	Asp	Glu	Val
530					535					540					

-310-

```

Arg Lys Leu Ala Glu Arg Thr Gln Lys Ser Leu Ser Glu Ile Glu Ala
545                               550                               555                               560
Asn Ile Asn Ile Leu Val Gln Ser Ile Ser Asp Thr Ser Glu Ser Ile
                               565                               570                               575
Lys Asn Gln Val Lys Glu Val Glu Glu Ile Asn Ala Ser Ile Glu Ala
                               580                               585                               590
Leu Arg Ser Val Thr Glu Gly Asn Leu Lys Ile Ala Ser Asp Ser Leu
595                               600                               605
Glu Ile Ser Gln Glu Ile Asp Lys Val Ser Asn Asp Ile Leu Glu Asp
610                               615                               620
Val Asn Lys Lys Gln Phe
625                               630

```

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1758 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 8...1702
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

```

GAGATAA ATG ATG TTT TCT TCA ATG TTT GCT TCG TTG GGG ACT CGT ATC      49
Met Met Phe Ser Ser Met Phe Ala Ser Leu Gly Thr Arg Ile
      1              5              10

ATG CTG GTC GTG TTA GCC GCT CTT TTA GGT TTA GGG GGG CTT TTT ATT      97
Met Leu Val Val Leu Ala Ala Leu Leu Gly Leu Gly Gly Leu Phe Ile
15              20              25              30

GGT TTT GTA AAG GTT ATG CAA AAA GAT GTG TTA GCG CAA CTC ATG GAG      145
Gly Phe Val Lys Val Met Gln Lys Asp Val Leu Ala Gln Leu Met Glu
      35              40              45

CAT TTA GAA ACC GGG CAA TAC AAA AAG CGT GAA AAA ACG CTC GCT TAC      193
His Leu Glu Thr Gly Gln Tyr Lys Lys Arg Glu Lys Thr Leu Ala Tyr
      50              55              60

ATG ACA AAA ATT ATT GAA CAG GGC ATT CAT GAG TAT TAC AAA AAT TTT      241
Met Thr Lys Ile Ile Glu Gln Gly Ile His Glu Tyr Tyr Lys Asn Phe
      65              70              75

GAC AAT GCT ACT GCA AGA AAA ATG GCG TTA GAT TAT TTC AAA CGC ATC      289
Asp Asn Ala Thr Ala Arg Lys Met Ala Leu Asp Tyr Phe Lys Arg Ile
      80              85              90

```

-311-

AAC GAC GAT AAG GGC ATG ATT TAT ATG GTG GTG GTG GAT AAA AAC GCG	337
Asn Asp Asp Lys Gly Met Ile Tyr Met Val Val Val Asp Lys Asn Gly	
95 100 105 110	
GTG GTA TTG TTT GAT CCG GTC AAT CCT AAA ACC GTA GNC CAA TCA GGG	385
Val Val Leu Phe Asp Pro Val Asn Pro Lys Thr Val Xaa Gln Ser Gly	
115 120 125	
CTT GAC GCT CAG AGC GTT GAT GGG GTG TAT TAT GTT AGG GGG TAT TTG	433
Leu Asp Ala Gln Ser Val Asp Gly Val Tyr Tyr Val Arg Gly Tyr Leu	
130 135 140	
GAG GCG GCC AAA AAA GGG GGA GGC TAC ACT TAT TAT AAA ATG CCT AAA	481
Glu Ala Ala Lys Lys Gly Gly Gly Tyr Thr Tyr Tyr Lys Met Pro Lys	
145 150 155	
TAC GAT GGA GGC GTA CCG GAG AAA AAA TTC GCC TAC TCG CAT TAT GAT	529
Tyr Asp Gly Gly Val Pro Glu Lys Lys Phe Ala Tyr Ser His Tyr Asp	
160 165 170	
GAA GTT TCT CAA ATG GTG ATC GCA ACG ACT TCC TAT TAC ACT GAC ATT	577
Glu Val Ser Gln Met Val Ile Ala Thr Thr Ser Tyr Tyr Thr Asp Ile	
175 180 185 190	
AAC ACA GAA AAT AAA GCG ATC AAA GAA GGC GTG AAT AAG GTT TTT GAT	625
Asn Thr Glu Asn Lys Ala Ile Lys Glu Gly Val Asn Lys Val Phe Asp	
195 200 205	
GAA AAC ACC ACG AAA TTA TTC CTT TGG ATA CTG ACA GCG ACG ATA GCG	673
Glu Asn Thr Thr Lys Leu Phe Leu Trp Ile Leu Thr Ala Thr Ile Ala	
210 215 220	
CTA GTG GTT TTG ACG CTC ATA TAC GCT AAA TTA AGG ATC GTG AAA CGC	721
Leu Val Val Leu Thr Leu Ile Tyr Ala Lys Leu Arg Ile Val Lys Arg	
225 230 235	
ATT GAT GAA CTG GTC CTT AAA ATC AAC GCT TTT AGC CGT GGG GAT AAG	769
Ile Asp Glu Leu Val Leu Lys Ile Asn Ala Phe Ser Arg Gly Asp Lys	
240 245 250	
GAT TTG AGA GCC AAA ATT GAT GTG GGT GAT CGC AAC GAT GAA ATC TCG	817
Asp Leu Arg Ala Lys Ile Asp Val Gly Asp Arg Asn Asp Glu Ile Ser	
255 260 265 270	
CAA GTG GGC CGT GGG ATC AAT TTG TTT GTG GAA AAC GCC CGC TTG ATT	865
Gln Val Gly Arg Gly Ile Asn Leu Phe Val Glu Asn Ala Arg Leu Ile	
275 280 285	
ATG GAA GAG ATT AAA GGG ATT TCC ACC CTC AAT AAA ACT TCA ATG GAT	913
Met Glu Glu Ile Lys Gly Ile Ser Thr Leu Asn Lys Thr Ser Met Asp	
290 295 300	
AAA TTA GTC CAA ATC ACG CAA GAA ACC CAA AAG AGC ATG AAA GAT TCC	961
Lys Leu Val Gln Ile Thr Gln Glu Thr Gln Lys Ser Met Lys Asp Ser	
305 310 315	

-312-

TCA ACC ACC CTA AAT TCC GTG AAA AAT AAA GCC ACT GAT ATA GCG AGC Ser Thr Thr Leu Asn Ser Val Lys Asn Lys Ala Thr Asp Ile Ala Ser 320 325 330	1009
ATG ATG AAT GCT TCC ATA GAG CAA TCT CAA GGG TTA AGG AAG CGT TTG Met Met Asn Ala Ser Ile Glu Gln Ser Gln Gly Leu Arg Lys Arg Leu 335 340 345 350	1057
ATT GAA ACG CAA GGG CTG GTC AAA GAG AGC AAG GAT GCG ATC GGG GAT Ile Glu Thr Gln Gly Leu Val Lys Glu Ser Lys Asp Ala Ile Gly Asp 355 360 365	1105
TTA TTT TCT CAA ATC ACA GAG AGC GCG CAC ACT GAA GAG GAA CTC TCT Leu Phe Ser Gln Ile Thr Glu Ser Ala His Thr Glu Glu Glu Leu Ser 370 375 380	1153
AGC AAA GTG GAG CAG CTA AGC CGT AAC GCT GAT GAT GTC AAA TCC ATT Ser Lys Val Glu Gln Leu Ser Arg Asn Ala Asp Asp Val Lys Ser Ile 385 390 395	1201
CTG GAT ATT ATC AAT GAT ATT GCC GAT CAA ACG AAT TTA TTA GCC CTA Leu Asp Ile Ile Asn Asp Ile Ala Asp Gln Thr Asn Leu Leu Ala Leu 400 405 410	1249
AAC GCT GCT ATT GAA GCC GCA AGG GCT GGC GAG CAT GGC AGA GGC TTT Asn Ala Ala Ile Glu Ala Ala Arg Ala Gly Glu His Gly Arg Gly Phe 415 420 425 430	1297
GCG GTG GTG GCT GAT GAA GTT AGG AAT TTA GCC GGG CGC ACT CAA AAG Ala Val Val Ala Asp Glu Val Arg Asn Leu Ala Gly Arg Thr Gln Lys 435 440 445	1345
TCT TTA GCC GAA ATC AAT TCC ACT ATC ATG GTG ATT GTC CAA GAA ATC Ser Leu Ala Glu Ile Asn Ser Thr Ile Met Val Ile Val Gln Glu Ile 450 455 460	1393
AAT GCC GTG AGT TCG CAA ATG AAT CTC AAT TCG CAA AAA ATG GAG CGT Asn Ala Val Ser Ser Gln Met Asn Leu Asn Ser Gln Lys Met Glu Arg 465 470 475	1441
TTG AGC GAT ATG AGT AAA AGC GTG CAA GAA ACT TAC GAA AAA ATG AGT Leu Ser Asp Met Ser Lys Ser Val Gln Glu Thr Tyr Glu Lys Met Ser 480 485 490	1489
TCT AAT TTA AGC TCA GTC GTG TCA GAC AGC AAT CAA AGC ATG GAC GAT Ser Asn Leu Ser Ser Val Val Ser Asp Ser Asn Gln Ser Met Asp Asp 495 500 505 510	1537
TAC GCC AAA TCC GGA CAC CAA ATT GAA GTT ATG GTA AGC GAT TTT GCA Tyr Ala Lys Ser Gly His Gln Ile Glu Val Met Val Ser Asp Phe Ala 515 520 525	1585
GAG GTG GAA AAA GTG GCT TCT AAG ACT TTA GCG GAT TCT TCA GAT ATT Glu Val Glu Lys Val Ala Ser Lys Thr Leu Ala Asp Ser Ser Asp Ile 530 535 540	1633

TTA AAC ATC GCT ACG CAT GTG AGT GGA ACG ACC ATG AAT TTA GAC AAA 1681
Leu Asn Ile Ala Thr His Val Ser Gly Thr Thr Met Asn Leu Asp Lys
545 550 555

CAA GTG AAT TTG TTT AAA ACT TAATCAGGGG GAGTTTATTA AAAAAGGGTT GGAT 1736
Gln Val Asn Leu Phe Lys Thr
560 565

(2) INFORMATION FOR SEO ID NO:162:

(A) LENGTH: 565 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

BNSDOCID: <WO 9821225A1.1 >

[illegible]

(A) LENGTH: 686 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- BNSDOCID: <WO 9821225A1_1_>

-315-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

TATAAGGTTG CTCTC ATG AAA AAA CCC TAT AGG AAG ATT TCT GAT TAT GCG	51
Met Lys Lys Pro Tyr Arg Lys Ile Ser Asp Tyr Ala	
1 5 10	
ATC GTG GGT GGT TTG AGC GCG TTA GTG ATG GTG AGC ATT GTG GGG TGT	99
Ile Val Gly Gly Leu Ser Ala Leu Val Met Val Ser Ile Val Gly Cys	
15 20 25	
AAG AGC AAT GCT GAT GAC AAA CCA AAA GAG CAA AGC TCT TTA AGT CAA	147
Lys Ser Asn Ala Asp Asp Lys Pro Lys Glu Gln Ser Ser Leu Ser Gln	
30 35 40	
AGC GTT CAA AAA GGC GCG TTT GTG ATT TTA GAA GAG CAA AAG GAT AAA	195
Ser Val Gln Lys Gly Ala Phe Val Ile Leu Glu Glu Gln Lys Asp Lys	
45 50 55 60	
TCT TAC AAG GTT GTT GAA GAA TAC CCC AGC TCA AGA ACC CAC ATT ATA	243
Ser Tyr Lys Val Val Glu Glu Tyr Pro Ser Ser Arg Thr His Ile Ile	
65 70 75	
GTG CGC GAT TTG CAA GGC AAT GAA CGC GTG TTA AGC AAT GAA GAG ATT	291
Val Arg Asp Leu Gln Gly Asn Glu Arg Val Leu Ser Asn Glu Glu Ile	
80 85 90	
CAA AAG CTC ATC AAA GAA GAA GAA GCT AAA ATT GAT AAC GGC ACG AGC	339
Gln Lys Leu Ile Lys Glu Glu Glu Ala Lys Ile Asp Asn Gly Thr Ser	
95 100 105	
AAG CTT GTC CAG CCT AAT AAT GGA GGG AGT AAT GAA GGC TCA GGC TTT	387
Lys Leu Val Gln Pro Asn Asn Gly Gly Ser Asn Glu Gly Ser Gly Phe	
110 115 120	
GGC TTG GGG AGC GCG ATT TTA GGG AGC GCG GCG GGG GCG ATT TTA GGG	435
Gly Leu Gly Ser Ala Ile Leu Gly Ser Ala Ala Gly Ala Ile Leu Gly	
125 130 135 140	
AGT TAT ATT GGT AAT AAG CTT TTC AAT AAC CCT AAT TAC CAG CAA AAC	483
Ser Tyr Ile Gly Asn Lys Leu Phe Asn Asn Pro Asn Tyr Gln Gln Asn	
145 150 155	
GCC CAA CGG ACC TAC AAA TCC CCA CAA GCT TAC CAA CGC TCT CAA AAT	531
Ala Gln Arg Thr Tyr Lys Ser Pro Gln Ala Tyr Gln Arg Ser Gln Asn	
160 165 170	
TCC TTT TCT AAA AGT GCG CCC AGT GCT TCA AGC ATG GGC GGA GCG AGT	579
Ser Phe Ser Lys Ser Ala Pro Ser Ala Ser Ser Met Gly Gly Ala Ser	
175 180 185	
AAG GGA CAG AGC GGG TTT TTT GGC TCT AGT AGG CCT ACT AGT TCA CCG	627
Lys Gly Gln Ser Gly Phe Phe Gly Ser Ser Arg Pro Thr Ser Ser Pro	
190 195 200	

GCG GTA AGC TCT GGG ACA AGG GGC TTT AAC TCA TAATTTAATT GATTCAAGGC 680
 Ala Val Ser Ser Gly Thr Arg Gly Phe Asn Ser
 205 210 215

686

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8748 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-317-

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 16...8694

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

AGAGGGTAGC ATTTA ATG AAA AAG TTT AAA AAG AAA CCA AAA AGT ATC AAA	51
Met Lys Lys Phe Lys Lys Lys Pro Lys Ser Ile Lys	
1 5 10	
CGA TCG CAT CAA AAT CAA AAA ACA ATC TTA AAG CGT CCT TTA TGG CTT	99
Arg Ser His Gln Asn Gln Lys Thr Ile Leu Lys Arg Pro Leu Trp Leu	
15 20 25	
ATG CCT TTA CTC ATC AGC GGG TTT GCT AGT GGG GTG TAT GCG AAT AAT	147
Met Pro Leu Leu Ile Ser Gly Phe Ala Ser Gly Val Tyr Ala Asn Asn	
30 35 40	
CTG TGG GAT TTG TTA AAC CCA AAA GTG GGG GGT GAG TAT GTG CAT TGG	195
Leu Trp Asp Leu Leu Asn Pro Lys Val Gly Gly Glu Tyr Val His Trp	
45 50 55 60	
GTT AAG GGC AGT CAG TAT TGT GCA TGG TGG GAA TTT GCT GGG TGT TTA	243
Val Lys Gly Ser Gln Tyr Cys Ala Trp Trp Glu Phe Ala Gly Cys Leu	
65 70 75	
AAG AAT GTA TGG GGG GCA AAT CAT AAA GGC TAT GAT GCT GGA AAC GCC	291
Lys Asn Val Trp Gly Ala Asn His Lys Gly Tyr Asp Ala Gly Asn Ala	
80 85 90	
GCT AAC TAT TTG TCT TCT CAA AAC TAT CAA GCT ATT TCG GTG GGT AGT	339
Ala Asn Tyr Leu Ser Ser Gln Asn Tyr Gln Ala Ile Ser Val Gly Ser	
95 100 105	
GGG AAT GAA ACG GGG ACT TAT AGT TTA AGC GGT TTT ACC AAT TAT GTT	387
Gly Asn Glu Thr Gly Thr Tyr Ser Leu Ser Gly Phe Thr Asn Tyr Val	
110 115 120	
GGG GGC AAT CTC ACG ATC AAT CTA GGC AAT AGC GTT GTT TTA GAT TTA	435
Gly Gly Asn Leu Thr Ile Asn Leu Gly Asn Ser Val Val Leu Asp Leu	
125 130 135 140	
AGC GGT TCT AAT AGT TTC ACT TCG TAT CAA GGT TAT AAT CAA GGC AAA	483
Ser Gly Ser Asn Ser Phe Thr Ser Tyr Gln Gly Tyr Asn Gln Gly Lys	
145 150 155	
GAT GAT GTA ACA TTT ACG GTT GGC GCA ATC AAT TTA AAC GGC ACT TTA	531
Asp Asp Val Thr Phe Thr Val Gly Ala Ile Asn Leu Asn Gly Thr Leu	
160 165 170	

-318-

GAA GTG GGT AAT CGT GTG GGA TCG GGA GCT GGC ACG CAC ACC GGC ACA	579
Glu Val Gly Asn Arg Val Gly Ser Gly Ala Gly Thr His Thr Gly Thr	
175 180 185	
GCC ACT TTA AAC TTG AAC GCT AAT AAG GTC AAT ATC AAT TCC AAT ATC	627
Ala Thr Leu Asn Leu Asn Ala Asn Lys Val Asn Ile Asn Ser Asn Ile	
190 195 200	
AAC GCG TAT AAA ACT TCG CAA GTG AAT ATA GGC AAC GCT AAC AGC GTT	675
Asn Ala Tyr Lys Thr Ser Gln Val Asn Ile Gly Asn Ala Asn Ser Val	
205 210 215 220	
ATT ACC ATT GGT TCG GTT TCT TTG AGT GGG GAT GTT TGC AGT TCT TTA	723
Ile Thr Ile Gly Ser Val Ser Leu Ser Gly Asp Val Cys Ser Ser Leu	
225 230 235	
GCT AGC GTT GGG ATA GGG GCT AAT TGC TCC ACT TCT GGG CCT AGC TAT	771
Ala Ser Val Gly Ile Gly Ala Asn Cys Ser Thr Ser Gly Pro Ser Tyr	
240 245 250	
TCT TTT AAA GGG ACG ACT AAC GCT ACT AAC ACG GCG TTT AGT AAT GCA	819
Ser Phe Lys Gly Thr Thr Asn Ala Thr Asn Thr Ala Phe Ser Asn Ala	
255 260 265	
AGC GGC AGT TTC ACT TTT GAA GAG AAC GCC ACT TTT AGC GGG GCG AAA	867
Ser Gly Ser Phe Thr Phe Glu Glu Asn Ala Thr Phe Ser Gly Ala Lys	
270 275 280	
TGG AAT GGG GGG ACT TAT ACC TTT AAT AAA GAG TTT AGC GCT ACC AAT	915
Trp Asn Gly Gly Thr Tyr Thr Phe Asn Lys Glu Phe Ser Ala Thr Asn	
285 290 295 300	
AAC ACC GCC TTT AGT AGC GGT AGT TTT AAT TTT AAA GGT GTA AGC TCT	963
Asn Thr Ala Phe Ser Ser Gly Ser Phe Asn Phe Lys Gly Val Ser Ser	
305 310 315	
TTT AAT GGT ACT TCG TTT AGT AAC GCT TCT TAT ACT TTT GAC AAT CAA	1011
Phe Asn Gly Thr Ser Phe Ser Asn Ala Ser Tyr Thr Phe Asp Asn Gln	
320 325 330	
GCC ACT TTC CAA AAC AGC TCC TTT AAT GGG GGG ACT TTT ACT TTT AAT	1059
Ala Thr Phe Gln Asn Ser Ser Phe Asn Gly Gly Thr Phe Thr Phe Asn	
335 340 345	
AAC CAA ACT AAT CCA ACT AAC AAC GCT CAG CAC CCC CAA ATT CAA AAC	1107
Asn Gln Thr Asn Pro Thr Asn Asn Ala Gln His Pro Gln Ile Gln Asn	
350 355 360	
AGC TCT TTT AGT GGT AAC GCT ACC ACT CTT AAG GGC TTT GTG AAT TTC	1155
Ser Ser Phe Ser Gly Asn Ala Thr Thr Leu Lys Gly Phe Val Asn Phe	
365 370 375 380	
CAG CAA GCC TTT AAC AAT TCA AAC CAC CAA CTA ACG ATC CAA AAC GCT	1203
Gln Gln Ala Phe Asn Asn Ser Asn His Gln Leu Thr Ile Gln Asn Ala	
385 390 395	

-319-

TCC TTT AAT AAC GCC AAT TTT AAC AAT ACC GGT AAA ATC ACT ATA GAA	1251
Ser Phe Asn Asn Ala Thr Phe Asn Asn Thr Gly Lys Ile Thr Ile Glu	
400 405 410	
AAA GAT GCG AGT TTT AAT AAC ACG ACA TTC AAC ACT TCT GTT GAT ACA	1299
Lys Asp Ala Ser Phe Asn Asn Thr Thr Phe Asn Thr Ser Val Asp Thr	
415 420 425	
AAC AAC ATG AGT GTT ACC GGT GGC GTT ACT TTA AGC GGT AAA AAT GAC	1347
Asn Asn Met Ser Val Thr Gly Gly Val Thr Leu Ser Gly Lys Asn Asp	
430 435 440	
TTG AAA AAT GGC TCA ACC CTT GAT TTT GGG AGT TCT AAA ATC ACT CTC	1395
Leu Lys Asn Gly Ser Thr Leu Asp Phe Gly Ser Ser Lys Ile Thr Leu	
445 450 455 460	
GCT CAA GGG ACG ACT TTC AAC CTC ACA AGT TTA GGC AGT GAG AAG AGC	1443
Ala Gln Gly Thr Thr Phe Asn Leu Thr Ser Leu Gly Ser Glu Lys Ser	
465 470 475	
GTA ACG ATT TTA AAT TCT AGC GGT GGG ATC ACT TAT AGT AAC CTT TTA	1491
Val Thr Ile Leu Asn Ser Ser Gly Gly Ile Thr Tyr Ser Asn Leu Leu	
480 485 490	
AAC CAT GCA ATC AAC GGC TTG ACA AGT GCC TTA AAA ACG AAC GAA AGC	1539
Asn His Ala Ile Asn Gly Leu Thr Ser Ala Leu Lys Thr Asn Glu Ser	
495 500 505	
CTT TCA AAT CCG CAA AGT TTC GCT CAA GGT TTG TGG GAT ATA ATC ACT	1587
Leu Ser Asn Pro Gln Ser Phe Ala Gln Gly Leu Trp Asp Ile Ile Thr	
510 515 520	
TAC AAT GGG GTT ACC GGG CAG CTT TTG AAT GAA AAC GCT GCA ACA TCT	1635
Tyr Asn Gly Val Thr Gly Gln Leu Leu Asn Glu Asn Ala Ala Thr Ser	
525 530 535 540	
AAA CCC ACT GAC TCT TCG CCC TCT AAA TCC TCT ACA AAC TCT ACG CAA	1683
Lys Pro Thr Asp Ser Ser Pro Ser Lys Ser Ser Thr Asn Ser Thr Gln	
545 550 555	
GTC TAT CAA GTG GGT TAC AAA ATA GGG GAT ACT ATC TAC AAA CTG CAA	1731
Val Tyr Gln Val Gly Tyr Lys Ile Gly Asp Thr Ile Tyr Lys Leu Gln	
560 565 570	
GAA ACT TTC AGC CAC AAT TCC ATT ATT ATT CAG GCT TTA GAG AGC GGG	1779
Glu Thr Phe Ser His Asn Ser Ile Ile Ile Gln Ala Leu Glu Ser Gly	
575 580 585	
ACT TAC ACG CCA CCC CCT GTC ATT AAC GGC TCC AAA TTT GAC TTA TCC	1827
Thr Tyr Thr Pro Pro Pro Val Ile Asn Gly Ser Lys Phe Asp Leu Ser	
590 595 600	
GCT TCA AAT TAT ATC AAT GCT GAC ATG CCT TGG TAT GAC CAT AAA TAT	1875
Ala Ser Asn Tyr Ile Asn Ala Asp Met Pro Trp Tyr Asp His Lys Tyr	
605 610 615 620	

- 320 -

TAC ATC CCT AAA TCC CAA AAT TTT ACA GAG AGC GGG ACT TAT TAC TTG	1923
Tyr Ile Pro Lys Ser Gln Asn Phe Thr Glu Ser Gly Thr Tyr Tyr Leu	
625 630 635	
CCG AGC GTC CAA ATA TGG GGG AGC TAC ACT AAC TCG TTT AAA CAA ACT	1971
Pro Ser Val Gln Ile Trp Gly Ser Tyr Thr Asn Ser Phe Lys Gln Thr	
640 645 650	
TTT AGC GCA AAT GGT AGT AAT CTG GTG ATT GGG TAT AAC TCA ACA TGG	2019
Phe Ser Ala Asn Gly Ser Asn Leu Val Ile Gly Tyr Asn Ser Thr Trp	
655 660 665	
ACT GAT CAT AAT GTC TCT TCT AGC GGC ACG GTG TCT TTT GGG GAC ACT	2067
Thr Asp His Asn Val Ser Ser Ser Gly Thr Val Ser Phe Gly Asp Thr	
670 675 680	
TCA GGG AGC GCT CTT AAT GGG CAT TGC GGA CCT TGG CCG TAT TAC CAA	2115
Ser Gly Ser Ala Leu Asn Gly His Cys Gly Pro Trp Pro Tyr Tyr Gln	
685 690 695 700	
TGC ACA GGC ACG ACT AAC GGC ACT TAT AGC GCC TAT CAT GTG TAT ATC	2163
Cys Thr Gly Thr Thr Asn Gly Thr Tyr Ser Ala Tyr His Val Tyr Ile	
705 710 715	
ACA GCG AAT CTG CGT TCT GGC AAT CGT ATA GGC ACC GGT GGG GCA GCT	2211
Thr Ala Asn Leu Arg Ser Gly Asn Arg Ile Gly Thr Gly Gly Ala Ala	
720 725 730	
AAT CTA ATC TTT AAT GGG GTA GAT AGT ATC AAT ATC GCT AAC GCT ACC	2259
Asn Leu Ile Phe Asn Gly Val Asp Ser Ile Asn Ile Ala Asn Ala Thr	
735 740 745	
ATC ACG CAA CAT AAC GCC GGA ATC TAT TCA AGC TCT ATG ACT TTT TCC	2307
Ile Thr Gln His Asn Ala Gly Ile Tyr Ser Ser Ser Met Thr Phe Ser	
750 755 760	
ACG CAA AGC ATG GAT AAT TCG CAG AAT TTG AAT GGT CTA AAT TCT AAC	2355
Thr Gln Ser Met Asp Asn Ser Gln Asn Leu Asn Gly Leu Asn Ser Asn	
765 770 775 780	
GGC AAA CTT TCG GTG TAT GGC ACC ACT TTC ACT AAC GAA GCT AAA GAT	2403
Gly Lys Leu Ser Val Tyr Gly Thr Thr Phe Thr Asn Glu Ala Lys Asp	
785 790 795	
GGG AAA TTC ATT TTC AAT GCA GGG CAA GCG GTT TTT GAA AAC ACC AAC	2451
Gly Lys Phe Ile Phe Asn Ala Gly Gln Ala Val Phe Glu Asn Thr Asn	
800 805 810	
TTT AAT GGA GGG AGT TAC CAA TTC AGC GGC GAT AGC TTG AAT TTT TCA	2499
Phe Asn Gly Gly Ser Tyr Gln Phe Ser Gly Asp Ser Leu Asn Phe Ser	
815 820 825	
AAC AAC AAC CAG TTC AAT AGC GGT TCG TTT GAA ATT AGC GCA AAA AAC	2547
Asn Asn Asn Gln Phe Asn Ser Gly Ser Phe Glu Ile Ser Ala Lys Asn	
830 835 840	

-321-

GCT TCG TTC AAT AAC GCT AAC TTT AAC AAC AGC GCT TCT TTT AAT TTC	2595
Ala Ser Phe Asn Asn Ala Asn Phe Asn Asn Ser Ala Ser Phe Asn Phe	
845 850 855 860	
AAT AAT TCT AAC GCG ACC ACT TCG TTT GTG GGG GAT TTC ACT AAC GCT	2643
Asn Asn Ser Asn Ala Thr Thr Ser Phe Val Gly Asp Phe Thr Asn Ala	
865 870 875	
AAT TCA AAT TTG CAA ATC GCC GGG AAC GCT GTT TTT GGG AAC TCT ACT	2691
Asn Ser Asn Leu Gln Ile Ala Gly Asn Ala Val Phe Gly Asn Ser Thr	
880 885 890	
AAT GGC TCT CAA AAT ACC GCT AAT TTT AAT AAT ACC GGC TCT GTT AAT	2739
Asn Gly Ser Gln Asn Thr Ala Asn Phe Asn Asn Thr Gly Ser Val Asn	
895 900 905	
ATT TCA GGG AAT GCA ACC TTT GAT AAT GTG GTG TTT AAT GGC CCT ACG	2787
Ile Ser Gly Asn Ala Thr Phe Asp Asn Val Val Phe Asn Gly Pro Thr	
910 915 920	
AAC ACG AGC GTG AAA GGG CAG GTT ACT TTA AAT AAC ATC ACT TTA AAA	2835
Asn Thr Ser Val Lys Gly Gln Val Thr Leu Asn Asn Ile Thr Leu Lys	
925 930 935 940	
AAC CTG AAC GCC CCT TTG TCT TTT GGC GAT GGG ACG ATT ACT TTT AAC	2883
Asn Leu Asn Ala Pro Leu Ser Phe Gly Asp Gly Thr Ile Thr Phe Asn	
945 950 955	
GCT CAT TCG GTG ATT AAT ATT GCT GAA TCT ATC ACT AAT GGC AAC CCT	2931
Ala His Ser Val Ile Asn Ile Ala Glu Ser Ile Thr Asn Gly Asn Pro	
960 965 970	
ATC ACT CTT GTA AGC TCT TCT AAA GAA ATT GAA TAC AAC AAC GCT TTC	2979
Ile Thr Leu Val Ser Ser Ser Lys Glu Ile Glu Tyr Asn Asn Ala Phe	
975 980 985	
AGT AAA AAT CTA TGG CAG CTC ATC AAC TAC CAA GGG CAT GGG GCA AGC	3027
Ser Lys Asn Leu Trp Gln Leu Ile Asn Tyr Gln Gly His Gly Ala Ser	
990 995 1000	
AGT GAA AAG CTC GTC TCT AGC GCG GGT AAT GGC GTT TAT GAT GTG GTG	3075
Ser Glu Lys Leu Val Ser Ser Ala Gly Asn Gly Val Tyr Asp Val Val	
1005 1010 1015 1020	
TAT TCT TTC AAT AAC CAA ACC TAC AAT TTC CAA GAG GTT TTT TCA CAA	3123
Tyr Ser Phe Asn Asn Gln Thr Tyr Asn Phe Gln Glu Val Phe Ser Gln	
1025 1030 1035	
AAC AGC ATT TCT ATC CGG CGT TTG GGC GTT AAC ATG GTG TTT GAT TAT	3171
Asn Ser Ile Ser Ile Arg Arg Leu Gly Val Asn Met Val Phe Asp Tyr	
1040 1045 1050	
GTG GAT ATG GAA AAA TCG GAT CAT TTA TAT TAT CAA AAC GCT CTC GGT	3219
Val Asp Met Glu Lys Ser Asp His Leu Tyr Tyr Gln Asn Ala Leu Gly	
1055 1060 1065	

-322-

TTT ATG ACC TAC ATG CCT AAT AGC TAT AAC AAT AAT TTA GGG AAT GCA	3267
Phe Met Thr Tyr Met Pro Asn Ser Tyr Asn Asn Asn Leu Gly Asn Ala	
1070 1075 1080	
AAC AAC ACC ATT TAC TAT TAC GAC AAG AGC ATT GAT TTT TAT GCG AGC	3315
Asn Asn Thr Ile Tyr Tyr Tyr Asp Lys Ser Ile Asp Phe Tyr Ala Ser	
1085 1090 1095 1100	
GGG AAA ACT CTA TTC ACT AAA GCG GAA TTT TCT CAA ACA TTC ACC GGG	3363
Gly Lys Thr Leu Phe Thr Lys Ala Glu Phe Ser Gln Thr Phe Thr Gly	
1105 1110 1115	
CAA AAC AGC GCG ATC GTT TTT GGG GCT AAA AGC ATA TGG ACG AGC TTA	3411
Gln Asn Ser Ala Ile Val Phe Gly Ala Lys Ser Ile Trp Thr Ser Leu	
1120 1125 1130	
AGC GAT GCA CCG CAG TCT AAC ACC ATC ATT CGC TTT GGG GAC AAT AAG	3459
Ser Asp Ala Pro Gln Ser Asn Thr Ile Ile Arg Phe Gly Asp Asn Lys	
1135 1140 1145	
GGA GCA GGG AGT AAT GAT GCG AGC GGG CAT TGC TGG AAT TTG CAA TGC	3507
Gly Ala Gly Ser Asn Asp Ala Ser Gly His Cys Trp Asn Leu Gln Cys	
1150 1155 1160	
ATA GGC TTT ATT ACA GGG CAT TAT GAA GCG CAA AAG ATT TAC ATC ACC	3555
Ile Gly Phe Ile Thr Gly His Tyr Glu Ala Gln Lys Ile Tyr Ile Thr	
1165 1170 1175 1180	
GGT AGC ATT GAA AGC GGG AAT CGC ATT TCT AGC GGT GGG GGC GCG AGC	3603
Gly Ser Ile Glu Ser Gly Asn Arg Ile Ser Ser Gly Gly Gly Ala Ser	
1185 1190 1195	
CTT AAT TTT AAC GGG CTT CAA GGC ATT CTT TTA ACG AAC GCG ACT TTG	3651
Leu Asn Phe Asn Gly Leu Gln Gly Ile Leu Leu Thr Asn Ala Thr Leu	
1200 1205 1210	
TAT AAC CGC GCC GCT GGC ACG CAA AGC TCG TCT ATG AAT TTT ATC TCT	3699
Tyr Asn Arg Ala Ala Gly Thr Gln Ser Ser Ser Met Asn Phe Ile Ser	
1215 1220 1225	
AAC AGC GCG AAC ATT CAG GCT CAA AAC TCC TAT TTT ATA GAC GAT ACC	3747
Asn Ser Ala Asn Ile Gln Ala Gln Asn Ser Tyr Phe Ile Asp Asp Thr	
1230 1235 1240	
GCA CAA AAT GGC GGT AAC CCT AAT TTC AGT TTC AAC GCT TTG AAT CTG	3795
Ala Gln Asn Gly Gly Asn Pro Asn Phe Ser Phe Asn Ala Leu Asn Leu	
1245 1250 1255 1260	
GAT TTT TCT AAC AGC TCT TTT AGA GGC TAT GTG GGG AAA ACG CAA TCT	3843
Asp Phe Ser Asn Ser Ser Phe Arg Gly Tyr Val Gly Lys Thr Gln Ser	
1265 1270 1275	
GTT TTT AAA TTC AAT GCC AAG AAT GCG ATC AGT TTC ACC AAC AGC ACG	3891
Val Phe Lys Phe Asn Ala Lys Asn Ala Ile Ser Phe Thr Asn Ser Thr	
1280 1285 1290	

-323-

AAT TTA AGC TCT GGT TTG TAT CAA ATG CAA GCT AAA AGC GTG TTG TTT	3939
Asn Leu Ser Ser Gly Leu Tyr Gln Met Gln Ala Lys Ser Val Leu Phe	
1295 1300 1305	
GAC AAT TCC AAT TTA AGC GTT TCA GTG GGG ACA AGC AGT ATT AAA GCC	3987
Asp Asn Ser Asn Leu Ser Val Ser Val Gly Thr Ser Ser Ile Lys Ala	
1310 1315 1320	
AAT GCG ATC AAT CTT TCT CAA AAT GCC TCT ATT AAT GCG AGC AAC CAT	4035
Asn Ala Ile Asn Leu Ser Gln Asn Ala Ser Ile Asn Ala Ser Asn His	
1325 1330 1335 1340	
TCA ACC TTA GAA CTT CAA GGC GAT TTG AAT GTG AAC GAC ACC AGC TCG	4083
Ser Thr Leu Glu Leu Gln Gly Asp Leu Asn Val Asn Asp Thr Ser Ser	
1345 1350 1355	
CTC AAC CTC AAC CAA AGC ACG ATT AAT GTT TCC AAT AAC GCC ACG ATC	4131
Leu Asn Leu Asn Gln Ser Thr Ile Asn Val Ser Asn Asn Ala Thr Ile	
1360 1365 1370	
AAC GAT TAT GCG AGC TTG ATT GCG AGT AAT GGC TCT CAC CTT AAT TTT	4179
Asn Asp Tyr Ala Ser Leu Ile Ala Ser Asn Gly Ser His Leu Asn Phe	
1375 1380 1385	
AAC GGG GCG GTT AAT TTC AAT TCA GCG AAT ATT ACT ACG AGT TTG AAT	4227
Asn Gly Ala Val Asn Phe Asn Ser Ala Asn Ile Thr Thr Ser Leu Asn	
1390 1395 1400	
AAT TCC TCT ATC GTG TTT AAG GGG GCG GTC TCT TTA GGA GGG CAG TTT	4275
Asn Ser Ser Ile Val Phe Lys Gly Ala Val Ser Leu Gly Gly Gln Phe	
1405 1410 1415 1420	
AAT TTA AGC AAT AAC TCT TCT TTA GAT TTC CAA GGC TCT AGC GCT ATC	4323
Asn Leu Ser Asn Asn Ser Ser Leu Asp Phe Gln Gly Ser Ser Ala Ile	
1425 1430 1435	
ACC TCT AAC ACG GCG TTT AAT TTC TAT GAT AAC GCT TTT TCT CAA AGC	4371
Thr Ser Asn Thr Ala Phe Asn Phe Tyr Asp Asn Ala Phe Ser Gln Ser	
1440 1445 1450	
CCC ATC ACT TTC CAT CAA GCC CTT GAC ATT AAA GCG CCC TTA AGT TTG	4419
Pro Ile Thr Phe His Gln Ala Leu Asp Ile Lys Ala Pro Leu Ser Leu	
1455 1460 1465	
GGA GGC AAC CTT TTA AAC CCT AAC AAC AGC AGC GTG CTG GAT TTA AAA	4467
Gly Gly Asn Leu Leu Asn Pro Asn Asn Ser Ser Val Leu Asp Leu Lys	
1470 1475 1480	
AAC AGC CAG CTT GTT TTT GGC GAT CAA GGG AGT TTG AAT ATC GCT AAC	4515
Asn Ser Gln Leu Val Phe Gly Asp Gln Gly Ser Leu Asn Ile Ala Asn	
1485 1490 1495 1500	
ATT GAT TTA CTA AGC GAT CTA AAT GAT AAT AAA AAT CGT GTG TAT AAC	4563
Ile Asp Leu Leu Ser Asp Leu Asn Asp Asn Lys Asn Arg Val Tyr Asn	
1505 1510 1515	

-324-

ATC ATT CAA GCG GAC ATG AAT AGT AAT TGG TAT GAG CGT ATC AGC TTC Ile Ile Gln Ala Asp Met Asn Ser Asn Trp Tyr Glu Arg Ile Ser Phe 1520 1525 1530	4611
TTT GGC ATG CAC ATC AAT GAC GGG ATT TAT GAT GCT AAA AAC CAA ACT Phe Gly Met His Ile Asn Asp Gly Ile Tyr Asp Ala Lys Asn Gln Thr 1535 1540 1545	4659
TAT AGT TTC ACT AAC CCC CTT AAT AAC GCC CTA AAA ATC ACC GAG AGC Tyr Ser Phe Thr Asn Pro Leu Asn Asn Ala Leu Lys Ile Thr Glu Ser 1550 1555 1560	4707
TTT AAA GAC AAC CAA CTA AGC GTT ACG CTC TCT CAA ATC CCG GGT ATT Phe Lys Asp Asn Gln Leu Ser Val Thr Leu Ser Gln Ile Pro Gly Ile 1565 1570 1575 1580	4755
AAA AAC ACG CTC TAT AAC ATT GGC TCT GAA ATT TTT AAC TAC CAA AAA Lys Asn Thr Leu Tyr Asn Ile Gly Ser Glu Ile Phe Asn Tyr Gln Lys 1585 1590 1595	4803
GTT TAT AAC AAC GCT AAT GGC GTG TAT TCT TAT AGC GAT GAT GCA CAA Val Tyr Asn Asn Ala Asn Gly Val Tyr Ser Tyr Ser Asp Asp Ala Gln 1600 1605 1610	4851
GGC GTG TTT TAT CTC ACA AGC AAC GTG AAA GGC TAT TAC AAC CCT AAC Gly Val Phe Tyr Leu Thr Ser Asn Val Lys Gly Tyr Tyr Asn Pro Asn 1615 1620 1625	4899
CAA TCC TAT CAA GCC AGC GGC AGT AAC AAC ACC ACG AAA AAT AAT AAT Gln Ser Tyr Gln Ala Ser Gly Ser Asn Asn Thr Thr Lys Asn Asn Asn 1630 1635 1640	4947
CTA ACC TCT GAA TCT TCT ATC ATC TCG CAA ACC TAT AAC GCG CAA GGC Leu Thr Ser Glu Ser Ser Ile Ile Ser Gln Thr Tyr Asn Ala Gln Gly 1645 1650 1655 1660	4995
AAC CCT ATT AGC GCG TTG CAC ATC TAT AAC AAG GGC TAT AAT TTC AAC Asn Pro Ile Ser Ala Leu His Ile Tyr Asn Lys Gly Tyr Asn Phe Asn 1665 1670 1675	5043
AAT ATC AAA GCG TTA GGG CAA ATG GCT CTC AAA CTC TAC CCT GAA ATC Asn Ile Lys Ala Leu Gly Gln Met Ala Leu Lys Leu Tyr Pro Glu Ile 1680 1685 1690	5091
AAA AAG GTA TTA GGG AAT GAT TTT TCG CCC TCA AGT TTG AAC GCT TTA Lys Lys Val Leu Gly Asn Asp Phe Ser Pro Ser Ser Leu Asn Ala Leu 1695 1700 1705	5139
AAC TCT AAT GCG CTA AAC CAA CTT ACC AAA CTC ATC ACG CCT AAC GAC Asn Ser Asn Ala Leu Asn Gln Leu Thr Lys Leu Ile Thr Pro Asn Asp 1710 1715 1720	5187
TGG AAA AAC ATT AAC GAG TTG ATT GAT AAC GCA AAC AAT TCG GTG GTG Trp Lys Asn Ile Asn Glu Leu Ile Asp Asn Ala Asn Asn Ser Val Val 1725 1730 1735 1740	5235

-325-

CAA AAT TTC AAT AAC GGC ACT TTG ATT GTG GGA GCG ACT CAA ATA GGG Gln Asn Phe Asn Asn Gly Thr Leu Ile Val Gly Ala Thr Gln Ile Gly 1745 1750 1755	5283
CAA ACA GAC ACC AAT AGC GCG GTT GTT TTT GGG GGC TTG GGC TAT CAA Gln Thr Asp Thr Asn Ser Ala Val Val Phe Gly Gly Leu Gly Tyr Gln 1760 1765 1770	5331
ACA CCT TGT GAT TAT ACT GAT ATT GTG TGC CAA AAA TTT AGA GGC ACT Thr Pro Cys Asp Tyr Thr Asp Ile Val Cys Gln Lys Phe Arg Gly Thr 1775 1780 1785	5379
TAT TTA GGA CAG CTT TTA GAG TCC AGC TCG GCT GAT TTG GGC TAT ATT Tyr Leu Gly Gln Leu Leu Glu Ser Ser Ser Ala Asp Leu Gly Tyr Ile 1790 1795 1800	5427
GAC ACG ACT TTT AAC GCT AAA GAA ATT TAT CTT ACC GGC ACT TTA GGG Asp Thr Thr Phe Asn Ala Lys Glu Ile Tyr Leu Thr Gly Thr Leu Gly 1805 1810 1815 1820	5475
AGC GGG AAC GCA TGG GGG ACT GGG GGG AGC GCG AGC GTA ACT TTT AAC Ser Gly Asn Ala Trp Gly Thr Gly Gly Ser Ala Ser Val Thr Phe Asn 1825 1830 1835	5523
AGC CAA ACT TCG CTC ATT CTC AAT CAG GCT AAT ATC GTA AGC TCG CAA Ser Gln Thr Ser Leu Ile Leu Asn Gln Ala Asn Ile Val Ser Ser Gln 1840 1845 1850	5571
ACC GAT GGG ATC TTT AGC ATG CTG GGT CAA GAG GGT ATT AAT AAG GTT Thr Asp Gly Ile Phe Ser Met Leu Gly Gln Glu Gly Ile Asn Lys Val 1855 1860 1865	5619
TTC AAT CAA GCC GGG CTC GCT AAT ATT TTG GGC GAA GTG GCG GTG CAA Phe Asn Gln Ala Gly Leu Ala Asn Ile Leu Gly Glu Val Ala Val Gln 1870 1875 1880	5667
TCC ATC AAC AAA GCC GGG GGA TTA GGG AAT TTG ATA GTA AAT ACG CTA Ser Ile Asn Lys Ala Gly Gly Leu Gly Asn Leu Ile Val Asn Thr Leu 1885 1890 1895 1900	5715
GGG AGT AAT AGC GTG ATT GGG GGG TAT TTA ACG CCT GAA CAA AAA AAT Gly Ser Asn Ser Val Ile Gly Gly Tyr Leu Thr Pro Glu Gln Lys Asn 1905 1910 1915	5763
CAA ACC CTA AGC CAG CTT TTA GGG CAG AAT AAC TTT GAT AAT CTC ATG Gln Thr Leu Ser Gln Leu Leu Gly Gln Asn Asn Phe Asp Asn Leu Met 1920 1925 1930	5811
AAC GAT AGC GGT TTG AAT ACG GCG ATT AAG GAT TTG ATC AGA CAA AAA Asn Asp Ser Gly Leu Asn Thr Ala Ile Lys Asp Leu Ile Arg Gln Lys 1935 1940 1945	5859
TTA GGC TTT TGG ACC GGG CTA GTG GGG GGA TTA GCC GGA CTA GGG GGC Leu Gly Phe Trp Thr Gly Leu Val Gly Gly Leu Ala Gly Leu Gly Gly 1950 1955 1960	5907

-326-

ATT GAT TTG CAA AAC CCT GAA AAG CTT ATA GGC AGC ATG TCA ATC AAT Ile Asp Leu Gln Asn Pro Glu Lys Leu Ile Gly Ser Met Ser Ile Asn 1965 1970 1975 1980	5955
GAT TTA TTG AGT AAA AAA GGG TTG TTC AAT CAG ATC ACC GGC TTT ATT Asp Leu Leu Ser Lys Lys Gly Leu Phe Asn Gln Ile Thr Gly Phe Ile 1985 1990 1995	6003
TCC GCT AAC GAT ATA GGG CAA GTC ATA AGC GTA ATG TTG CAA GAT ATT Ser Ala Asn Asp Ile Gly Gln Val Ile Ser Val Met Leu Gln Asp Ile 2000 2005 2010	6051
GTC AAA CCG AGC AAC GCT TTA AAA AAC GAT GTA GCG GCT TTA GGC AAG Val Lys Pro Ser Asn Ala Leu Lys Asn Asp Val Ala Ala Leu Gly Lys 2015 2020 2025	6099
CAA ATG ATT GGC GAA TTT TTA GGC CAA GAC ACG CTC AAT TCT TTA GAA Gln Met Ile Gly Glu Phe Leu Gly Gln Asp Thr Leu Asn Ser Leu Glu 2030 2035 2040	6147
AGC TTG TTG CAA AAC CAG CAG ATT AAA AGC GTT TTA GAC AAA GTC CTA Ser Leu Leu Gln Asn Gln Gln Ile Lys Ser Val Leu Asp Lys Val Leu 2045 2050 2055 2060	6195
GCG GCT AAA GGT TTA GGG CCT ATT TAT GAA CAA GGC TTG GGG GAT TTG Ala Ala Lys Gly Leu Gly Pro Ile Tyr Glu Gln Gly Leu Gly Asp Leu 2065 2070 2075	6243
ATA CCT AAT CTT GGT AAA AAA GGG CTT TTC GCT CCT TAT GGC TTG AGT Ile Pro Asn Leu Gly Lys Lys Gly Leu Phe Ala Pro Tyr Gly Leu Ser 2080 2085 2090	6291
CAA GTG TGG CAA AAA GGG GAT TTT AGT TTC AAC GCA CAA GGC AAT GTT Gln Val Trp Gln Lys Gly Asp Phe Ser Phe Asn Ala Gln Gly Asn Val 2095 2100 2105	6339
TTT GTG CAA AAT TCC ACT TTC TCT AAC GCC AAT GGA GGC ACG CTC TCT Phe Val Gln Asn Ser Thr Phe Ser Asn Ala Asn Gly Gly Thr Leu Ser 2110 2115 2120	6387
TTT AAC GCA GGA AAT TCG CTC ATT TTT GCC GGA AAC AAT CAT ATT GCA Phe Asn Ala Gly Asn Ser Leu Ile Phe Ala Gly Asn Asn His Ile Ala 2125 2130 2135 2140	6435
TTC ACT AAC CAC GCT GGA ACT CTT CAA TTA TTG TCC GAT CAA GTT TCT Phe Thr Asn His Ala Gly Thr Leu Gln Leu Leu Ser Asp Gln Val Ser 2145 2150 2155	6483
AAC ATT AAC ATC ACC ACG CTT AAC GCT AGC AAC GGC CTT AAG ATT AAC Asn Ile Asn Ile Thr Thr Leu Asn Ala Ser Asn Gly Leu Lys Ile Asn 2160 2165 2170	6531
GCC GCT AAT AAC AAT GTT TCT GTG TCT CAA GGC AAT CTG TTT GTC AGC Ala Ala Asn Asn Asn Val Ser Val Ser Gln Gly Asn Leu Phe Val Ser 2175 2180 2185	6579

-327-

GCT AGC TGC GCG CAA CAA AGC GAT CCA ACT ACA GCT AAT ATT GCA AAC Ala Ser Cys Ala Gln Gln Ser Asp Pro Thr Thr Ala Asn Ile Ala Asn 2190 2195 2200	6627
CCT TGC GCG CTT AGC GCC CAA AGC ACG AAT GGC GCT TCT TCT AAT AAT Pro Cys Ala Leu Ser Ala Gln Ser Thr Asn Gly Ala Ser Ser Asn Asn 2205 2210 2215 2220	6675
GCG TCA AAT AAC GCG CCA ATC GCC TTG AGT AAT AAC GAT GAA AGC TTG Ala Ser Asn Asn Ala Pro Ile Ala Leu Ser Asn Asn Asp Glu Ser Leu 2225 2230 2235	6723
ATG GTT GCG GCG AAT GAT TTC AAT TTT TCA GGC AAT ATT TAC GCT AAT Met Val Ala Ala Asn Asp Phe Asn Phe Ser Gly Asn Ile Tyr Ala Asn 2240 2245 2250	6771
GGG GTG GTT GAT TTT TCA AAG ATT AAA GGC TCT GCA AAC ATT AAA AAC Gly Val Val Asp Phe Ser Lys Ile Lys Gly Ser Ala Asn Ile Lys Asn 2255 2260 2265	6819
CTG TAT CTT TAC AAT AAC GCT CAA TTC CAA GCC AAC AAT CTC ACT ATT Leu Tyr Leu Tyr Asn Asn Ala Gln Phe Gln Ala Asn Asn Leu Thr Ile 2270 2275 2280	6867
TCC AAT CAA GCG GTG TTA GAA AAA AAC GCC AGC TTT GTA ACG AAT AAT Ser Asn Gln Ala Val Leu Glu Lys Asn Ala Ser Phe Val Thr Asn Asn 2285 2290 2295 2300	6915
TTA AAC ATT CAA GGA GCG TTT AAC AAC AAC GCC ACG CAA AAA ATA GAG Leu Asn Ile Gln Gly Ala Phe Asn Asn Asn Ala Thr Gln Lys Ile Glu 2305 2310 2315	6963
GTG CTT CAA AAT TTA GTG ATC GCT TCA AAC GCT TCT TTA AGC ACC GGG Val Leu Gln Asn Leu Val Ile Ala Ser Asn Ala Ser Leu Ser Thr Gly 2320 2325 2330	7011
ATT TAT GGG TTA GAA GTA GGG GGG GCT TTG AAT AAT TCT GGA GCG ATC Ile Tyr Gly Leu Glu Val Gly Gly Ala Leu Asn Asn Ser Gly Ala Ile 2335 2340 2345	7059
CAT TTT AAT TTA GAA AAT ACC CAA ACG CCA ACG CCG CTC ATT CAA GCA His Phe Asn Leu Glu Asn Thr Gln Thr Pro Thr Pro Leu Ile Gln Ala 2350 2355 2360	7107
GAG GGG ATC ATT AAC CTC AAC ACC ACC CAA ACG CCT TTT ATG AAT GTC Glu Gly Ile Ile Asn Leu Asn Thr Thr Gln Thr Pro Phe Met Asn Val 2365 2370 2375 2380	7155
AAT AAC AGC ATG GCC AAT AAT ACG ACT TAC ACT TTA TTA AAA AGC AGC Asn Asn Ser Met Ala Asn Asn Thr Thr Tyr Thr Leu Leu Lys Ser Ser 2385 2390 2395	7203
CGT TAC ATT GAT TAC AAT ATC AAC CCC AAC AGC TTG CAA TCG TAT TTG Arg Tyr Ile Asp Tyr Asn Ile Asn Pro Asn Ser Leu Gln Ser Tyr Leu 2400 2405 2410	7251

-328-

AAT CTC TAC ACT TTA ATC AAT ATC AAC GGG AAC CAC ATA GAG GAA AAA Asn Leu Tyr Thr Leu Ile Asn Ile Asn Gly Asn His Ile Glu Glu Lys 2415 2420 2425	7299
AAC GGC GCA TTG ACT TAT TTG GGC CAA CGG GTT TTG TTG CAA GAT AAG Asn Gly Ala Leu Thr Tyr Leu Gly Gln Arg Val Leu Leu Gln Asp Lys 2430 2435 2440	7347
GGG TTA TTG TTA AGC GTA GCG CTG CCC AAC TCA AAC AAC GCT TCT CAA Gly Leu Leu Leu Ser Val Ala Leu Pro Asn Ser Asn Asn Ala Ser Gln 2445 2450 2455 2460	7395
AAC AAC ATT TTA AGC CTT TCT GTC CTT TAT AAC CAA GTT AAA ATG TCT Asn Asn Ile Leu Ser Leu Ser Val Leu Tyr Asn Gln Val Lys Met Ser 2465 2470 2475	7443
TGC GGC GAT AAA GCG ATG GAT TTT ACC CCC CCT ACC TTA CAA GAT TAC Cys Gly Asp Lys Ala Met Asp Phe Thr Pro Pro Thr Leu Gln Asp Tyr 2480 2485 2490	7491
ATT GTG GGC ATT CAA GGG CAA AGC GCG CTC AAT CAA ATT GAA GCT GTT Ile Val Gly Ile Gln Gly Gln Ser Ala Leu Asn Gln Ile Glu Ala Val 2495 2500 2505	7539
GGG GGC AAC GCT ATC AAG TGG CTT TCA ACA TTG ATG ATG GAG ACT AAA Gly Gly Asn Ala Ile Lys Trp Leu Ser Thr Leu Met Met Glu Thr Lys 2510 2515 2520	7587
GAA AAC CCG TTT TTT GCG CCG ATT TAT TTA AAA AAC CAC TCT TTG AAT Glu Asn Pro Phe Phe Ala Pro Ile Tyr Leu Lys Asn His Ser Leu Asn 2525 2530 2535 2540	7635
GAA ATC TTA GGC GTA ACA AAA GAT CTT CAA AAC ACC GCA AGC TTG ATT Glu Ile Leu Gly Val Thr Lys Asp Leu Gln Asn Thr Ala Ser Leu Ile 2545 2550 2555	7683
TCT AAC CCT AAT TTT AGA GAT AAC GCT ACC AAT CTT TTA GAA TTG GCG Ser Asn Pro Asn Phe Arg Asp Asn Ala Thr Asn Leu Leu Glu Leu Ala 2560 2565 2570	7731
AGT TAC ACC CAA CAA ACC AGC CGT TTA ACA AAA CTC TCT GAT TTT AGA Ser Tyr Thr Gln Gln Thr Ser Arg Leu Thr Lys Leu Ser Asp Phe Arg 2575 2580 2585	7779
TCT AGA GAG GGA GAG TCT GAT TTT TCT TTG TTA GAG CTT AAA AAC AAG Ser Arg Glu Gly Glu Ser Asp Phe Ser Leu Leu Glu Leu Lys Asn Lys 2590 2595 2600	7827
CGT TTT AGC GAT CCT AAT CCA GAG GTT TTT GTC AAA TAC TCT CAA CTT Arg Phe Ser Asp Pro Asn Pro Glu Val Phe Val Lys Tyr Ser Gln Leu 2605 2610 2615 2620	7875
AGC AAA CAC CCA AAT AAC CTT TGG GTT CAA GGG GTG GGA GGA GCG AGC Ser Lys His Pro Asn Asn Leu Trp Val Gln Gly Val Gly Gly Ala Ser 2625 2630 2635	7923

-329-

TTT ATT TCT GGG GGC AAT GGC ACG CTT TAT GGC TTG AAT GCG GGC TAT	7971
Phe Ile Ser Gly Gly Asn Gly Thr Leu Tyr Gly Leu Asn Ala Gly Tyr	
2640 2645 2650	
GAC AGG TTG GTT AAA AAT GTG ATC CTT GGG GGT TAT GTG GCT TAT GGC	8019
Asp Arg Leu Val Lys Asn Val Ile Leu Gly Gly Tyr Val Ala Tyr Gly	
2655 2660 2665	
TAT AGC GAC TTT AAT GGG AAC ATC ATG CAT TCT TTG GGT AAT AAT GTG	8067
Tyr Ser Asp Phe Asn Gly Asn Ile Met His Ser Leu Gly Asn Asn Val	
2670 2675 2680	
GAT GTG GGG ATG TAT GCG AGG GCT TTT TTA AAA AGG AAC GAA TTC ACT	8115
Asp Val Gly Met Tyr Ala Arg Ala Phe Leu Lys Arg Asn Glu Phe Thr	
2685 2690 2695 2700	
TTG AGC GCG AAT GAA ACT TAT GGA GGC AAT GCA ACT AGT ATC AAT TCT	8163
Leu Ser Ala Asn Glu Thr Tyr Gly Gly Asn Ala Thr Ser Ile Asn Ser	
2705 2710 2715	
TCT AAT TCT TTG CTC TCT GTG TTG AAC CAA CGC TAC AAC TAC AAC ACC	8211
Ser Asn Ser Leu Leu Ser Val Leu Asn Gln Arg Tyr Asn Tyr Asn Thr	
2720 2725 2730	
TGG ACA ACG AGC GTG AAC GGG AAT TAC GGC TAT GAT TTC ATG TTC AAA	8259
Trp Thr Thr Ser Val Asn Gly Asn Tyr Gly Tyr Asp Phe Met Phe Lys	
2735 2740 2745	
CAA AAA AGC GTG GTG CTA AAA CCT CAA GTG GGT TTG AGC TAT CAT TTC	8307
Gln Lys Ser Val Val Leu Lys Pro Gln Val Gly Leu Ser Tyr His Phe	
2750 2755 2760	
ATA GGT CTA AGT GGG ATG AAA GGC AAT GAT GCC GCT TAC AAA CAA TTC	8355
Ile Gly Leu Ser Gly Met Lys Gly Asn Asp Ala Ala Tyr Lys Gln Phe	
2765 2770 2775 2780	
CTC ATG CAT TCA AAC CCC TCT AAC GAA TCG GTT TTA ACG CTC AAC ATG	8403
Leu Met His Ser Asn Pro Ser Asn Glu Ser Val Leu Thr Leu Asn Met	
2785 2790 2795	
GGG TTG GAG AGC CGT AAA TAT TTT GGT AAA AAT TCC TAT TAT TTT GTA	8451
Gly Leu Glu Ser Arg Lys Tyr Phe Gly Lys Asn Ser Tyr Tyr Phe Val	
2800 2805 2810	
ACG GCG AGA CTA GGT AGG GAT CTT TTG ATC AAA TCT AAA GGC AGC AAT	8499
Thr Ala Arg Leu Gly Arg Asp Leu Leu Ile Lys Ser Lys Gly Ser Asn	
2815 2820 2825	
ACG GTG CGT TTT GTG GGC GAA AAC ACT TTA TTG TAT CGC AAG GGG GAA	8547
Thr Val Arg Phe Val Gly Glu Asn Thr Leu Leu Tyr Arg Lys Gly Glu	
2830 2835 2840	
GTT TTT AAC ACT TTT GCG AGC GTG ATT ACA GGG GGC GAA ATG CAT TTG	8595
Val Phe Asn Thr Phe Ala Ser Val Ile Thr Gly Gly Glu Met His Leu	
2845 2850 2855 2860	

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2893 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

BNSDOCID: <WO 9821225A1_1_>

-331-

Ile	Gly	Ala	Asn	Cys	Ser	Thr	Ser	Gly	Pro	Ser	Tyr	Ser	Phe	Lys	Gly
				245					250					255	
Thr	Thr	Asn	Ala	Thr	Asn	Thr	Ala	Phe	Ser	Asn	Ala	Ser	Gly	Ser	Phe
			260					265					270		
Thr	Phe	Glu	Glu	Asn	Ala	Thr	Phe	Ser	Gly	Ala	Lys	Trp	Asn	Gly	Gly
		275					280					285			
Thr	Tyr	Thr	Phe	Asn	Lys	Glu	Phe	Ser	Ala	Thr	Asn	Asn	Thr	Ala	Phe
	290				295						300				
Ser	Ser	Gly	Ser	Phe	Asn	Phe	Lys	Gly	Val	Ser	Ser	Phe	Asn	Gly	Thr
305					310					315				320	
Ser	Phe	Ser	Asn	Ala	Ser	Tyr	Thr	Phe	Asp	Asn	Gln	Ala	Thr	Phe	Gln
			325					330						335	
Asn	Ser	Ser	Phe	Asn	Gly	Gly	Thr	Phe	Thr	Phe	Asn	Asn	Gln	Thr	Asn
			340					345					350		
Pro	Thr	Asn	Asn	Ala	Gln	His	Pro	Gln	Ile	Gln	Asn	Ser	Ser	Phe	Ser
		355					360					365			
Gly	Asn	Ala	Thr	Thr	Leu	Lys	Gly	Phe	Val	Asn	Phe	Gln	Gln	Ala	Phe
	370					375					380				
Asn	Asn	Ser	Asn	His	Gln	Leu	Thr	Ile	Gln	Asn	Ala	Ser	Phe	Asn	Asn
385					390				395					400	
Ala	Thr	Phe	Asn	Asn	Thr	Gly	Lys	Ile	Thr	Ile	Glu	Lys	Asp	Ala	Ser
			405					410						415	
Phe	Asn	Asn	Thr	Thr	Phe	Asn	Thr	Ser	Val	Asp	Thr	Asn	Asn	Met	Ser
			420					425					430		
Val	Thr	Gly	Gly	Val	Thr	Leu	Ser	Gly	Lys	Asn	Asp	Leu	Lys	Asn	Gly
	435						440					445			
Ser	Thr	Leu	Asp	Phe	Gly	Ser	Ser	Lys	Ile	Thr	Leu	Ala	Gln	Gly	Thr
	450					455					460				
Thr	Phe	Asn	Leu	Thr	Ser	Leu	Gly	Ser	Glu	Lys	Ser	Val	Thr	Ile	Leu
465					470					475				480	
Asn	Ser	Ser	Gly	Gly	Ile	Thr	Tyr	Ser	Asn	Leu	Leu	Asn	His	Ala	Ile
			485					490						495	
Asn	Gly	Leu	Thr	Ser	Ala	Leu	Lys	Thr	Asn	Glu	Ser	Leu	Ser	Asn	Pro
		500						505					510		
Gln	Ser	Phe	Ala	Gln	Gly	Leu	Trp	Asp	Ile	Ile	Thr	Tyr	Asn	Gly	Val
	515						520					525			
Thr	Gly	Gln	Leu	Leu	Asn	Glu	Asn	Ala	Ala	Thr	Ser	Lys	Pro	Thr	Asp
	530					535					540				
Ser	Ser	Pro	Ser	Lys	Ser	Ser	Thr	Asn	Ser	Thr	Gln	Val	Tyr	Gln	Val
545					550				555					560	
Gly	Tyr	Lys	Ile	Gly	Asp	Thr	Ile	Tyr	Lys	Leu	Gln	Glu	Thr	Phe	Ser
			565					570						575	
His	Asn	Ser	Ile	Ile	Ile	Gln	Ala	Leu	Glu	Ser	Gly	Thr	Tyr	Thr	Pro
		580					585						590		
Pro	Pro	Val	Ile	Asn	Gly	Ser	Lys	Phe	Asp	Leu	Ser	Ala	Ser	Asn	Tyr
		595					600					605			
Ile	Asn	Ala	Asp	Met	Pro	Trp	Tyr	Asp	His	Lys	Tyr	Tyr	Ile	Pro	Lys
	610					615					620				
Ser	Gln	Asn	Phe	Thr	Glu	Ser	Gly	Thr	Tyr	Tyr	Leu	Pro	Ser	Val	Gln
625					630					635				640	
Ile	Trp	Gly	Ser	Tyr	Thr	Asn	Ser	Phe	Lys	Gln	Thr	Phe	Ser	Ala	Asn
			645					650						655	
Gly	Ser	Asn	Leu	Val	Ile	Gly	Tyr	Asn	Ser	Thr	Trp	Thr	Asp	His	Asn
		660					665					670			
Val	Ser	Ser	Ser	Gly	Thr	Val	Ser	Phe	Gly	Asp	Thr	Ser	Gly	Ser	Ala

-332-

	675					680					685				
Leu	Asn	Gly	His	Cys	Gly	Pro	Trp	Pro	Tyr	Tyr	Gln	Cys	Thr	Gly	Thr
	690					695					700				
Thr	Asn	Gly	Thr	Tyr	Ser	Ala	Tyr	His	Val	Tyr	Ile	Thr	Ala	Asn	Leu
705					710					715					720
Arg	Ser	Gly	Asn	Arg	Ile	Gly	Thr	Gly	Gly	Ala	Ala	Asn	Leu	Ile	Phe
				725					730					735	
Asn	Gly	Val	Asp	Ser	Ile	Asn	Ile	Ala	Asn	Ala	Thr	Ile	Thr	Gln	His
		740					745					750			
Asn	Ala	Gly	Ile	Tyr	Ser	Ser	Ser	Met	Thr	Phe	Ser	Thr	Gln	Ser	Met
	755						760				765				
Asp	Asn	Ser	Gln	Asn	Leu	Asn	Gly	Leu	Asn	Ser	Asn	Gly	Lys	Leu	Ser
	770					775					780				
Val	Tyr	Gly	Thr	Thr	Phe	Thr	Asn	Glu	Ala	Lys	Asp	Gly	Lys	Phe	Ile
785					790					795					800
Phe	Asn	Ala	Gly	Gln	Ala	Val	Phe	Glu	Asn	Thr	Asn	Phe	Asn	Gly	Gly
				805					810					815	
Ser	Tyr	Gln	Phe	Ser	Gly	Asp	Ser	Leu	Asn	Phe	Ser	Asn	Asn	Asn	Gln
			820					825					830		
Phe	Asn	Ser	Gly	Ser	Phe	Glu	Ile	Ser	Ala	Lys	Asn	Ala	Ser	Phe	Asn
	835						840					845			
Asn	Ala	Asn	Phe	Asn	Asn	Ser	Ala	Ser	Phe	Asn	Phe	Asn	Asn	Ser	Asn
	850					855					860				
Ala	Thr	Thr	Ser	Phe	Val	Gly	Asp	Phe	Thr	Asn	Ala	Asn	Ser	Asn	Leu
865					870					875					880
Gln	Ile	Ala	Gly	Asn	Ala	Val	Phe	Gly	Asn	Ser	Thr	Asn	Gly	Ser	Gln
				885					890					895	
Asn	Thr	Ala	Asn	Phe	Asn	Asn	Thr	Gly	Ser	Val	Asn	Ile	Ser	Gly	Asn
		900						905					910		
Ala	Thr	Phe	Asp	Asn	Val	Val	Phe	Asn	Gly	Pro	Thr	Asn	Thr	Ser	Val
	915							920					925		
Lys	Gly	Gln	Val	Thr	Leu	Asn	Asn	Ile	Thr	Leu	Lys	Asn	Leu	Asn	Ala
	930					935					940				
Pro	Leu	Ser	Phe	Gly	Asp	Gly	Thr	Ile	Thr	Phe	Asn	Ala	His	Ser	Val
945					950					955					960
Ile	Asn	Ile	Ala	Glu	Ser	Ile	Thr	Asn	Gly	Asn	Pro	Ile	Thr	Leu	Val
				965					970					975	
Ser	Ser	Ser	Lys	Glu	Ile	Glu	Tyr	Asn	Asn	Ala	Phe	Ser	Lys	Asn	Leu
		980						985					990		
Trp	Gln	Leu	Ile	Asn	Tyr	Gln	Gly	His	Gly	Ala	Ser	Ser	Glu	Lys	Leu
	995					1000						1005			
Val	Ser	Ser	Ala	Gly	Asn	Gly	Val	Tyr	Asp	Val	Val	Tyr	Ser	Phe	Asn
	1010					1015						1020			
Asn	Gln	Thr	Tyr	Asn	Phe	Gln	Glu	Val	Phe	Ser	Gln	Asn	Ser	Ile	Ser
025					1030					1035					1040
Ile	Arg	Arg	Leu	Gly	Val	Asn	Met	Val	Phe	Asp	Tyr	Val	Asp	Met	Glu
				1045					1050					1055	
Lys	Ser	Asp	His	Leu	Tyr	Tyr	Gln	Asn	Ala	Leu	Gly	Phe	Met	Thr	Tyr
		1060						1065				1070			
Met	Pro	Asn	Ser	Tyr	Asn	Asn	Asn	Leu	Gly	Asn	Ala	Asn	Asn	Thr	Ile
	1075					1080						1085			
Tyr	Tyr	Tyr	Asp	Lys	Ser	Ile	Asp	Phe	Tyr	Ala	Ser	Gly	Lys	Thr	Leu
	1090					1095						1100			
Phe	Thr	Lys	Ala	Glu	Phe	Ser	Gln	Thr	Phe	Thr	Gly	Gln	Asn	Ser	Ala
105					1110					1115					1120

-333-

Ile Val Phe Gly Ala Lys-Ser Ile Trp Thr Ser Leu Ser Asp Ala Pro
 1125 1130 1135
 Gln Ser Asn Thr Ile Ile Arg Phe Gly Asp Asn Lys Gly Ala Gly Ser
 1140 1145 1150
 Asn Asp Ala Ser Gly His Cys Trp Asn Leu Gln Cys Ile Gly Phe Ile
 1155 1160 1165
 Thr Gly His Tyr Glu Ala Gln Lys Ile Tyr Ile Thr Gly Ser Ile Glu
 1170 1175 1180
 Ser Gly Asn Arg Ile Ser Ser Gly Gly Gly Ala Ser Leu Asn Phe Asn
 185 1190 1195 1200
 Gly Leu Gln Gly Ile Leu Leu Thr Asn Ala Thr Leu Tyr Asn Arg Ala
 1205 1210 1215
 Ala Gly Thr Gln Ser Ser Ser Met Asn Phe Ile Ser Asn Ser Ala Asn
 1220 1225 1230
 Ile Gln Ala Gln Asn Ser Tyr Phe Ile Asp Asp Thr Ala Gln Asn Gly
 1235 1240 1245
 Gly Asn Pro Asn Phe Ser Phe Asn Ala Leu Asn Leu Asp Phe Ser Asn
 1250 1255 1260
 Ser Ser Phe Arg Gly Tyr Val Gly Lys Thr Gln Ser Val Phe Lys Phe
 265 1270 1275 1280
 Asn Ala Lys Asn Ala Ile Ser Phe Thr Asn Ser Thr Asn Leu Ser Ser
 1285 1290 1295
 Gly Leu Tyr Gln Met Gln Ala Lys Ser Val Leu Phe Asp Asn Ser Asn
 1300 1305 1310
 Leu Ser Val Ser Val Gly Thr Ser Ser Ile Lys Ala Asn Ala Ile Asn
 1315 1320 1325
 Leu Ser Gln Asn Ala Ser Ile Asn Ala Ser Asn His Ser Thr Leu Glu
 1330 1335 1340
 Leu Gln Gly Asp Leu Asn Val Asn Asp Thr Ser Ser Leu Asn Leu Asn
 345 1350 1355 1360
 Gln Ser Thr Ile Asn Val Ser Asn Asn Ala Thr Ile Asn Asp Tyr Ala
 1365 1370 1375
 Ser Leu Ile Ala Ser Asn Gly Ser His Leu Asn Phe Asn Gly Ala Val
 1380 1385 1390
 Asn Phe Asn Ser Ala Asn Ile Thr Thr Ser Leu Asn Asn Ser Ser Ile
 1395 1400 1405
 Val Phe Lys Gly Ala Val Ser Leu Gly Gly Gln Phe Asn Leu Ser Asn
 1410 1415 1420
 Asn Ser Ser Leu Asp Phe Gln Gly Ser Ser Ala Ile Thr Ser Asn Thr
 425 1430 1435 1440
 Ala Phe Asn Phe Tyr Asp Asn Ala Phe Ser Gln Ser Pro Ile Thr Phe
 1445 1450 1455
 His Gln Ala Leu Asp Ile Lys Ala Pro Leu Ser Leu Gly Gly Asn Leu
 1460 1465 1470
 Leu Asn Pro Asn Asn Ser Ser Val Leu Asp Leu Lys Asn Ser Gln Leu
 1475 1480 1485
 Val Phe Gly Asp Gln Gly Ser Leu Asn Ile Ala Asn Ile Asp Leu Leu
 1490 1495 1500
 Ser Asp Leu Asn Asp Asn Lys Asn Arg Val Tyr Asn Ile Ile Gln Ala
 505 1510 1515 1520
 Asp Met Asn Ser Asn Trp Tyr Glu Arg Ile Ser Phe Phe Gly Met His
 1525 1530 1535
 Ile Asn Asp Gly Ile Tyr Asp Ala Lys Asn Gln Thr Tyr Ser Phe Thr
 1540 1545 1550
 Asn Pro Leu Asn Asn Ala Leu Lys Ile Thr Glu Ser Phe Lys Asp Asn

-334-

1555 1560 1565
 Gln Leu Ser Val Thr Leu Ser Gln Ile Pro Gly Ile Lys Asn Thr Leu
 1570 1575 1580
 Tyr Asn Ile Gly Ser Glu Ile Phe Asn Tyr Gln Lys Val Tyr Asn Asn
 585 1590 1595 1600
 Ala Asn Gly Val Tyr Ser Tyr Ser Asp Asp Ala Gln Gly Val Phe Tyr
 1605 1610 1615
 Leu Thr Ser Asn Val Lys Gly Tyr Tyr Asn Pro Asn Gln Ser Tyr Gln
 1620 1625 1630
 Ala Ser Gly Ser Asn Asn Thr Thr Lys Asn Asn Asn Leu Thr Ser Glu
 1635 1640 1645
 Ser Ser Ile Ile Ser Gln Thr Tyr Asn Ala Gln Gly Asn Pro Ile Ser
 1650 1655 1660
 Ala Leu His Ile Tyr Asn Lys Gly Tyr Asn Phe Asn Asn Ile Lys Ala
 665 1670 1675 1680
 Leu Gly Gln Met Ala Leu Lys Leu Tyr Pro Glu Ile Lys Lys Val Leu
 1685 1690 1695
 Gly Asn Asp Phe Ser Pro Ser Ser Leu Asn Ala Leu Asn Ser Asn Ala
 1700 1705 1710
 Leu Asn Gln Leu Thr Lys Leu Ile Thr Pro Asn Asp Trp Lys Asn Ile
 1715 1720 1725
 Asn Glu Leu Ile Asp Asn Ala Asn Asn Ser Val Val Gln Asn Phe Asn
 1730 1735 1740
 Asn Gly Thr Leu Ile Val Gly Ala Thr Gln Ile Gly Gln Thr Asp Thr
 745 1750 1755 1760
 Asn Ser Ala Val Val Phe Gly Gly Leu Gly Tyr Gln Thr Pro Cys Asp
 1765 1770 1775
 Tyr Thr Asp Ile Val Cys Gln Lys Phe Arg Gly Thr Tyr Leu Gly Gln
 1780 1785 1790
 Leu Leu Glu Ser Ser Ser Ala Asp Leu Gly Tyr Ile Asp Thr Thr Phe
 1795 1800 1805
 Asn Ala Lys Glu Ile Tyr Leu Thr Gly Thr Leu Gly Ser Gly Asn Ala
 1810 1815 1820
 Trp Gly Thr Gly Gly Ser Ala Ser Val Thr Phe Asn Ser Gln Thr Ser
 825 1830 1835 1840
 Leu Ile Leu Asn Gln Ala Asn Ile Val Ser Ser Gln Thr Asp Gly Ile
 1845 1850 1855
 Phe Ser Met Leu Gly Gln Glu Gly Ile Asn Lys Val Phe Asn Gln Ala
 1860 1865 1870
 Gly Leu Ala Asn Ile Leu Gly Glu Val Ala Val Gln Ser Ile Asn Lys
 1875 1880 1885
 Ala Gly Gly Leu Gly Asn Leu Ile Val Asn Thr Leu Gly Ser Asn Ser
 1890 1895 1900
 Val Ile Gly Gly Tyr Leu Thr Pro Glu Gln Lys Asn Gln Thr Leu Ser
 905 1910 1915 1920
 Gln Leu Leu Gly Gln Asn Asn Phe Asp Asn Leu Met Asn Asp Ser Gly
 1925 1930 1935
 Leu Asn Thr Ala Ile Lys Asp Leu Ile Arg Gln Lys Leu Gly Phe Trp
 1940 1945 1950
 Thr Gly Leu Val Gly Gly Leu Ala Gly Leu Gly Gly Ile Asp Leu Gln
 1955 1960 1965
 Asn Pro Glu Lys Leu Ile Gly Ser Met Ser Ile Asn Asp Leu Leu Ser
 1970 1975 1980
 Lys Lys Gly Leu Phe Asn Gln Ile Thr Gly Phe Ile Ser Ala Asn Asp
 985 1990 1995 2000

- 335 -

Ile	Gly	Gln	Val	Ile	Ser	Val	Met	Leu	Gln	Asp	Ile	Val	Lys	Pro	Ser
2005				2010				2015							
Asn	Ala	Leu	Lys	Asn	Asp	Val	Ala	Ala	Leu	Gly	Lys	Gln	Met	Ile	Gly
2020				2025				2030							
Glu	Phe	Leu	Gly	Gln	Asp	Thr	Leu	Asn	Ser	Leu	Glu	Ser	Leu	Leu	Gln
2035				2040				2045							
Asn	Gln	Gln	Ile	Lys	Ser	Val	Leu	Asp	Lys	Val	Leu	Ala	Ala	Lys	Gly
2050				2055				2060							
Leu	Gly	Pro	Ile	Tyr	Glu	Gln	Gly	Leu	Gly	Asp	Leu	Ile	Pro	Asn	Leu
065	2070				2075				2080						
Gly	Lys	Lys	Gly	Leu	Phe	Ala	Pro	Tyr	Gly	Leu	Ser	Gln	Val	Trp	Gln
2085				2090				2095							
Lys	Gly	Asp	Phe	Ser	Phe	Asn	Ala	Gln	Gly	Asn	Val	Phe	Val	Gln	Asn
2100				2105				2110							
Ser	Thr	Phe	Ser	Asn	Ala	Asn	Gly	Gly	Thr	Leu	Ser	Phe	Asn	Ala	Gly
2115				2120				2125							
Asn	Ser	Leu	Ile	Phe	Ala	Gly	Asn	Asn	His	Ile	Ala	Phe	Thr	Asn	His
2130				2135				2140							
Ala	Gly	Thr	Leu	Gln	Leu	Leu	Ser	Asp	Gln	Val	Ser	Asn	Ile	Asn	Ile
145	2150				2155				2160						
Thr	Thr	Leu	Asn	Ala	Ser	Asn	Gly	Leu	Lys	Ile	Asn	Ala	Ala	Asn	Asn
2165				2170				2175							
Asn	Val	Ser	Val	Ser	Gln	Gly	Asn	Leu	Phe	Val	Ser	Ala	Ser	Cys	Ala
2180				2185				2190							
Gln	Gln	Ser	Asp	Pro	Thr	Thr	Ala	Asn	Ile	Ala	Asn	Pro	Cys	Ala	Leu
2195				2200				2205							
Ser	Ala	Gln	Ser	Thr	Asn	Gly	Ala	Ser	Ser	Asn	Asn	Ala	Ser	Asn	Asn
2210				2215				2220							
Ala	Pro	Ile	Ala	Leu	Ser	Asn	Asn	Asp	Glu	Ser	Leu	Met	Val	Ala	Ala
225	2230				2235				2240						
Asn	Asp	Phe	Asn	Phe	Ser	Gly	Asn	Ile	Tyr	Ala	Asn	Gly	Val	Val	Asp
2245				2250				2255							
Phe	Ser	Lys	Ile	Lys	Gly	Ser	Ala	Asn	Ile	Lys	Asn	Leu	Tyr	Leu	Tyr
2260				2265				2270							
Asn	Asn	Ala	Gln	Phe	Gln	Ala	Asn	Asn	Leu	Thr	Ile	Ser	Asn	Gln	Ala
2275				2280				2285							
Val	Leu	Glu	Lys	Asn	Ala	Ser	Phe	Val	Thr	Asn	Asn	Leu	Asn	Ile	Gln
2290				2295				2300							
Gly	Ala	Phe	Asn	Asn	Asn	Ala	Thr	Gln	Lys	Ile	Glu	Val	Leu	Gln	Asn
305	2310				2315				2320						
Leu	Val	Ile	Ala	Ser	Asn	Ala	Ser	Leu	Ser	Thr	Gly	Ile	Tyr	Gly	Leu
2325				2330				2335							
Glu	Val	Gly	Gly	Ala	Leu	Asn	Asn	Ser	Gly	Ala	Ile	His	Phe	Asn	Leu
2340				2345				2350							
Glu	Asn	Thr	Gln	Thr	Pro	Thr	Pro	Leu	Ile	Gln	Ala	Glu	Gly	Ile	Ile
2355				2360				2365							
Asn	Leu	Asn	Thr	Thr	Gln	Thr	Pro	Phe	Met	Asn	Val	Asn	Asn	Ser	Met
2370				2375				2380							
Ala	Asn	Asn	Thr	Thr	Tyr	Thr	Leu	Leu	Lys	Ser	Ser	Arg	Tyr	Ile	Asp
385	2390				2395				2400						
Tyr	Asn	Ile	Asn	Pro	Asn	Ser	Leu	Gln	Ser	Tyr	Leu	Asn	Leu	Tyr	Thr
2405				2410				2415							
Leu	Ile	Asn	Ile	Asn	Gly	Asn	His	Ile	Glu	Glu	Lys	Asn	Gly	Ala	Leu
2420				2425				2430							
Thr	Tyr	Leu	Gly	Gln	Arg	Val	Leu	Leu	Gln	Asp	Lys	Gly	Leu	Leu	

-336-

2435	2440	2445
Ser Val Ala Leu Pro Asn Ser Asn Asn Ala Ser Gln Asn Asn Ile Leu		
2450	2455	2460
Ser Leu Ser Val Leu Tyr Asn Gln Val Lys Met Ser Cys Gly Asp Lys		
465	2470	2475
Ala Met Asp Phe Thr Pro Pro Thr Leu Gln Asp Tyr Ile Val Gly Ile		2480
	2485	2490
Gln Gly Gln Ser Ala Leu Asn Gln Ile Glu Ala Val Gly Gly Asn Ala		2495
2500	2505	2510
Ile Lys Trp Leu Ser Thr Leu Met Met Glu Thr Lys Glu Asn Pro Phe		
2515	2520	2525
Phe Ala Pro Ile Tyr Leu Lys Asn His Ser Leu Asn Glu Ile Leu Gly		
2530	2535	2540
Val Thr Lys Asp Leu Gln Asn Thr Ala Ser Leu Ile Ser Asn Pro Asn		
545	2550	2555
Phe Arg Asp Asn Ala Thr Asn Leu Leu Glu Leu Ala Ser Tyr Thr Gln		2560
	2565	2570
Gln Thr Ser Arg Leu Thr Lys Leu Ser Asp Phe Arg Ser Arg Glu Gly		2575
2580	2585	2590
Glu Ser Asp Phe Ser Leu Leu Glu Leu Lys Asn Lys Arg Phe Ser Asp		
2595	2600	2605
Pro Asn Pro Glu Val Phe Val Lys Tyr Ser Gln Leu Ser Lys His Pro		
2610	2615	2620
Asn Asn Leu Trp Val Gln Gly Val Gly Gly Ala Ser Phe Ile Ser Gly		
625	2630	2635
Gly Asn Gly Thr Leu Tyr Gly Leu Asn Ala Gly Tyr Asp Arg Leu Val		2640
	2645	2650
Lys Asn Val Ile Leu Gly Gly Tyr Val Ala Tyr Gly Tyr Ser Asp Phe		2655
2660	2665	2670
Asn Gly Asn Ile Met His Ser Leu Gly Asn Asn Val Asp Val Gly Met		
2675	2680	2685
Tyr Ala Arg Ala Phe Leu Lys Arg Asn Glu Phe Thr Leu Ser Ala Asn		
2690	2695	2700
Glu Thr Tyr Gly Gly Asn Ala Thr Ser Ile Asn Ser Ser Asn Ser Leu		
705	2710	2715
Leu Ser Val Leu Asn Gln Arg Tyr Asn Tyr Asn Thr Trp Thr Thr Ser		2720
	2725	2730
Val Asn Gly Asn Tyr Gly Tyr Asp Phe Met Phe Lys Gln Lys Ser Val		2735
2740	2745	2750
Val Leu Lys Pro Gln Val Gly Leu Ser Tyr His Phe Ile Gly Leu Ser		
2755	2760	2765
Gly Met Lys Gly Asn Asp Ala Ala Tyr Lys Gln Phe Leu Met His Ser		
2770	2775	2780
Asn Pro Ser Asn Glu Ser Val Leu Thr Leu Asn Met Gly Leu Glu Ser		
785	2790	2795
Arg Lys Tyr Phe Gly Lys Asn Ser Tyr Tyr Phe Val Thr Ala Arg Leu		2800
	2805	2810
Gly Arg Asp Leu Leu Ile Lys Ser Lys Gly Ser Asn Thr Val Arg Phe		2815
2820	2825	2830
Val Gly Glu Asn Thr Leu Leu Tyr Arg Lys Gly Glu Val Phe Asn Thr		
2835	2840	2845
Phe Ala Ser Val Ile Thr Gly Gly Glu Met His Leu Trp Arg Leu Val		
2850	2855	2860
Tyr Val Asn Ala Gly Val Gly Leu Lys Met Gly Leu Gln Tyr Gln Asp		
865	2870	2875
		2880

-337-

Ile Asn Ile Thr Gly Asn-Val Gly Met Arg Val Ala Phe
 2885 2890

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1376 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 13...1338
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

TGGTAGTTAA GA ATG GGT AAT CAT TTT TCT AAA TTA GGA TTT GTT TTA GCC	51
Met Gly Asn His Phe Ser Lys Leu Gly Phe Val Leu Ala	
1 5 10	
GCA TTA GGA AGC GCG ATA GGT TTA GGG CAT ATC TGG CGT TTC CCC TAC	99
Ala Leu Gly Ser Ala Ile Gly Leu Gly His Ile Trp Arg Phe Pro Tyr	
15 20 25	
ATG ACT GGG GTG AGT GGT GGG GGT GCT TTT GTT TTA TTG TTT TTA TTT	147
Met Thr Gly Val Ser Gly Gly Gly Ala Phe Val Leu Leu Phe Leu Phe	
30 35 40 45	
TTA TCT TTA AGC GTT GGC GCG GCG ATG TTT ATC GCT GAA ATG CTA TTA	195
Leu Ser Leu Ser Val Gly Ala Ala Met Phe Ile Ala Glu Met Leu Leu	
50 55 60	
GGA CAA AGC ACT CAA AAA AAT GTA ACA GAA GCT TTT AAA GAG CTT GAC	243
Gly Gln Ser Thr Gln Lys Asn Val Thr Glu Ala Phe Lys Glu Leu Asp	
65 70 75	
ATT AAC CCC AAA AAA CGC TGG AAA TAC GCA GGG CTT TTG CTT GTT TCT	291
Ile Asn Pro Lys Lys Arg Trp Lys Tyr Ala Gly Leu Leu Leu Val Ser	
80 85 90	
GGG CCA TTA ATA CTG ACT TTT TAC GGC ACG ATT TTA GGT TGG GTG CTT	339
Gly Pro Leu Ile Leu Thr Phe Tyr Gly Thr Ile Leu Gly Trp Val Leu	
95 100 105	
TAT TAT TTG GTG AGT GTT AGT TTT AAT TTG CCT AAC AAT ATC CAA GAA	387
Tyr Tyr Leu Val Ser Val Ser Phe Asn Leu Pro Asn Asn Ile Gln Glu	
110 115 120 125	

-338-

TCT GAA CAA ATT TTT ACT-CAA ACT TTG CAG TCT ATA GGG CTA CAA TCC Ser Glu Gln Ile Phe Thr Gln Thr Leu Gln Ser Ile Gly Leu Gln Ser 130 135 140	435
ATA GGG CTT TTT AGC GTT TTA TTG ATA ACC GGA TGG ATT GTT TCT AGG Ile Gly Leu Phe Ser Val Leu Leu Ile Thr Gly Trp Ile Val Ser Arg 145 150 155	483
GGG ATT AAA GAA GGC ATT GAA AAG CTC AAT TTG GTT TTA ATG CCC TTA Gly Ile Lys Glu Gly Ile Glu Lys Leu Asn Leu Val Leu Met Pro Leu 160 165 170	531
CTC TTT GCT ACT TTT TTT GGT TTG CTT TTC TAT GCG ATG AGC ATG GAT Leu Phe Ala Thr Phe Phe Gly Leu Leu Phe Tyr Ala Met Ser Met Asp 175 180 185	579
TCT TTT TCT AAA GCT TTT CAT TTC ATG TTT GAT TTC AAA CCA AAA GAT Ser Phe Ser Lys Ala Phe His Phe Met Phe Asp Phe Lys Pro Lys Asp 190 195 200 205	627
TTG ACC TCT CAA GTG TTC ACT TAT TCC TTG GGG CAG GTT TTC TTT TCC Leu Thr Ser Gln Val Phe Thr Tyr Ser Leu Gly Gln Val Phe Phe Ser 210 215 220	675
TTA AGC ATC GGT TTA GGG ATC AAT ATC ACT TAC GCT GCG GTT ACG GAT Leu Ser Ile Gly Leu Gly Ile Asn Ile Thr Tyr Ala Ala Val Thr Asp 225 230 235	723
AAA ACG CAG AAT TTG CTT AAA AGC ACT ATT TGG GTG GTT TTA TCA GGA Lys Thr Gln Asn Leu Leu Lys Ser Thr Ile Trp Val Val Leu Ser Gly 240 245 250	771
ATT CTA ATT TCT CTT GTG GCA GGA CTT ATG ATT TTC ACT TTT GTG TTT Ile Leu Ile Ser Leu Val Ala Gly Leu Met Ile Phe Thr Phe Val Phe 255 260 265	819
GAA TAT GGG GCG AAT GTC TCA CAA GGC ACA GGG TTA ATC TTC ACT TCT Glu Tyr Gly Ala Asn Val Ser Gln Gly Thr Gly Leu Ile Phe Thr Ser 270 275 280 285	867
TTA CCG GTG GTT TTT GGC CAA ATG GGA GCG ATA GGC ATT CTT GTT TCG Leu Pro Val Val Phe Gly Gln Met Gly Ala Ile Gly Ile Leu Val Ser 290 295 300	915
ATT CTT TTC TTG CTC GCG CTC GCT TTT GCT GGC ATC ACT TCT ACG GTG Ile Leu Phe Leu Leu Ala Leu Ala Phe Ala Gly Ile Thr Ser Thr Val 305 310 315	963
GCT TTA TTG GAG CCA AGC GTG ATG TAT CTT ACC GAA AGG TAT CAA TAC Ala Leu Leu Glu Pro Ser Val Met Tyr Leu Thr Glu Arg Tyr Gln Tyr 320 325 330	1011
TCT CGT TTT AAG GTT ACT TGG GGT CTT GTA GCA CTA ATT TTT GTG GTA Ser Arg Phe Lys Val Thr Trp Gly Leu Val Ala Leu Ile Phe Val Val 335 340 345	1059

-339-

```

GGC GTG GTG TTG ATT TTC-TCG CTC CAT AAG GAT TAT AAA GAT TAT CTC      1107
Gly Val Val Leu Ile Phe Ser Leu His Lys Asp Tyr Lys Asp Tyr Leu
350                      355                      360                      365

ACT TTC TTT GAA AAA AGT CTT TTT GAT TGG TTG GAT TTT GCA TCA AGC      1155
Thr Phe Phe Glu Lys Ser Leu Phe Asp Trp Leu Asp Phe Ala Ser Ser
                      370                      375                      380

ACC ATT ATC ATG CCT TTA GGC GGG ATG GCA ACC TTT ATT TTT ATG GGT      1203
Thr Ile Ile Met Pro Leu Gly Gly Met Ala Thr Phe Ile Phe Met Gly
                      385                      390                      395

TGG GTT TTG AAA AAA GAA AAA TTG CGT CTT TTG AGC GTG CAC TTT TTA      1251
Trp Val Leu Lys Lys Glu Lys Leu Arg Leu Leu Ser Val His Phe Leu
                      400                      405                      410

GGC CCT AAA TTG TTT GCA ACT TGG TAT TTC TTG CTT AAA TAT ATC ACC      1299
Gly Pro Lys Leu Phe Ala Thr Trp Tyr Phe Leu Leu Lys Tyr Ile Thr
                      415                      420                      425

CCT TTA ATT GTG TTT TCC ATT TGG TTG AGC AAG ATT TAT TAAAAATATTT GG  1350
Pro Leu Ile Val Phe Ser Ile Trp Leu Ser Lys Ile Tyr
430                      435                      440

CATGGGAAAA TTTTCTAAAT TAGGCT                                         1376

```

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 442 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

```

Met Gly Asn His Phe Ser Lys Leu Gly Phe Val Leu Ala Ala Leu Gly
 1                      5                      10                      15
Ser Ala Ile Gly Leu Gly His Ile Trp Arg Phe Pro Tyr Met Thr Gly
                      20                      25                      30
Val Ser Gly Gly Gly Ala Phe Val Leu Leu Phe Leu Phe Leu Ser Leu
                      35                      40                      45
Ser Val Gly Ala Ala Met Phe Ile Ala Glu Met Leu Leu Gly Gln Ser
                      50                      55                      60
Thr Gln Lys Asn Val Thr Glu Ala Phe Lys Glu Leu Asp Ile Asn Pro
65                      70                      75                      80
Lys Lys Arg Trp Lys Tyr Ala Gly Leu Leu Leu Val Ser Gly Pro Leu
                      85                      90                      95
Ile Leu Thr Phe Tyr Gly Thr Ile Leu Gly Trp Val Leu Tyr Tyr Leu
                      100                      105                      110
Val Ser Val Ser Phe Asn Leu Pro Asn Asn Ile Gln Glu Ser Glu Gln
                      115                      120                      125
Ile Phe Thr Gln Thr Leu Gln Ser Ile Gly Leu Gln Ser Ile Gly Leu
130                      135                      140

```

-340-

Phe	Ser	Val	Leu	Leu	Ile	Thr	Gly	Trp	Ile	Val	Ser	Arg	Gly	Ile	Lys
145					150					155					160
Glu	Gly	Ile	Glu	Lys	Leu	Asn	Leu	Val	Leu	Met	Pro	Leu	Leu	Phe	Ala
				165					170						175
Thr	Phe	Phe	Gly	Leu	Leu	Phe	Tyr	Ala	Met	Ser	Met	Asp	Ser	Phe	Ser
			180					185					190		
Lys	Ala	Phe	His	Phe	Met	Phe	Asp	Phe	Lys	Pro	Lys	Asp	Leu	Thr	Ser
		195					200					205			
Gln	Val	Phe	Thr	Tyr	Ser	Leu	Gly	Gln	Val	Phe	Phe	Ser	Leu	Ser	Ile
	210					215					220				
Gly	Leu	Gly	Ile	Asn	Ile	Thr	Tyr	Ala	Ala	Val	Thr	Asp	Lys	Thr	Gln
225					230					235					240
Asn	Leu	Leu	Lys	Ser	Thr	Ile	Trp	Val	Val	Leu	Ser	Gly	Ile	Leu	Ile
			245					250						255	
Ser	Leu	Val	Ala	Gly	Leu	Met	Ile	Phe	Thr	Phe	Val	Phe	Glu	Tyr	Gly
			260					265					270		
Ala	Asn	Val	Ser	Gln	Gly	Thr	Gly	Leu	Ile	Phe	Thr	Ser	Leu	Pro	Val
		275					280					285			
Val	Phe	Gly	Gln	Met	Gly	Ala	Ile	Gly	Ile	Leu	Val	Ser	Ile	Leu	Phe
	290					295					300				
Leu	Leu	Ala	Leu	Ala	Phe	Ala	Gly	Ile	Thr	Ser	Thr	Val	Ala	Leu	Leu
305					310					315					320
Glu	Pro	Ser	Val	Met	Tyr	Leu	Thr	Glu	Arg	Tyr	Gln	Tyr	Ser	Arg	Phe
			325					330						335	
Lys	Val	Thr	Trp	Gly	Leu	Val	Ala	Leu	Ile	Phe	Val	Val	Gly	Val	Val
			340					345					350		
Leu	Ile	Phe	Ser	Leu	His	Lys	Asp	Tyr	Lys	Asp	Tyr	Leu	Thr	Phe	Phe
		355					360					365			
Glu	Lys	Ser	Leu	Phe	Asp	Trp	Leu	Asp	Phe	Ala	Ser	Ser	Thr	Ile	Ile
	370					375					380				
Met	Pro	Leu	Gly	Gly	Met	Ala	Thr	Phe	Ile	Phe	Met	Gly	Trp	Val	Leu
385					390					395					400
Lys	Lys	Glu	Lys	Leu	Arg	Leu	Leu	Ser	Val	His	Phe	Leu	Gly	Pro	Lys
			405						410					415	
Leu	Phe	Ala	Thr	Trp	Tyr	Phe	Leu	Leu	Lys	Tyr	Ile	Thr	Pro	Leu	Ile
			420					425					430		
Val	Phe	Ser	Ile	Trp	Leu	Ser	Lys	Ile	Tyr						
			435				440								

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1392 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 22...1356
- (D) OTHER INFORMATION:

-341-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

TTTAAAAGGT ATTTTATAAC G ATG AAA ATT TTT GGG ACT GAT GGC GTG AGG	51
Met Lys Ile Phe Gly Thr Asp Gly Val Arg	
1 5 10	
GGT AAA GCA GGG GTG AAA CTC ACC CCC ATG TTT GTG ATG CGT TTA GGC	99
Gly Lys Ala Gly Val Lys Leu Thr Pro Met Phe Val Met Arg Leu Gly	
15 20 25	
ATT GCT GCC GGA TTG TAT TTT AAA AAA CAT TCT CAA ACG AAT AAA ATT	147
Ile Ala Ala Gly Leu Tyr Phe Lys Lys His Ser Gln Thr Asn Lys Ile	
30 35 40	
CTA ATC GGT AAA GAC ACC AGA AAA AGC GGC TAT ATG GTA GAA AAC GCT	195
Leu Ile Gly Lys Asp Thr Arg Lys Ser Gly Tyr Met Val Glu Asn Ala	
45 50 55	
TTA GTG AGC GCT CTA ACT TCC ATA GGC TAT AAT GTG ATT CAA ATA GGG	243
Leu Val Ser Ala Leu Thr Ser Ile Gly Tyr Asn Val Ile Gln Ile Gly	
60 65 70	
CCT ATG CCC ACC CCT GCG ATT GCG TTT TTA ACT GAA GAC ATG CGC TGT	291
Pro Met Pro Thr Pro Ala Ile Ala Phe Leu Thr Glu Asp Met Arg Cys	
75 80 85 90	
GAT GCG GGT ATT ATG ATA AGC GCG AGC CAC AAC CCT TTT GAA GAT AAT	339
Asp Ala Gly Ile Met Ile Ser Ala Ser His Asn Pro Phe Glu Asp Asn	
95 100 105	
GGC ATT AAG TTT TTC AAT TCT TAT GGC TAT AAG CTT AAA GAA GAA GAA	387
Gly Ile Lys Phe Phe Asn Ser Tyr Gly Tyr Lys Leu Lys Glu Glu Glu	
110 115 120	
GAA AAA GCG ATT GAA GAA ATC TTT CAT GAT GAA GAA TTA CTG CAT TCT	435
Glu Lys Ala Ile Glu Glu Ile Phe His Asp Glu Glu Leu Leu His Ser	
125 130 135	
AGC TAT AAA GTG GGT GAG AGC GTC GGT AGC GCT AAA AGG ATA GAC GAT	483
Ser Tyr Lys Val Gly Glu Ser Val Gly Ser Ala Lys Arg Ile Asp Asp	
140 145 150	
GTC ATA GGG CGC TAT ATT GCA CAT TTA AAA CAC TCT TTC CCC AAA CAT	531
Val Ile Gly Arg Tyr Ile Ala His Leu Lys His Ser Phe Pro Lys His	
155 160 165 170	
TTG AAT TTA CAG AGT TTA AGG ATC GTG CTA GAT ACG GCT AAT GGC GCG	579
Leu Asn Leu Gln Ser Leu Arg Ile Val Leu Asp Thr Ala Asn Gly Ala	
175 180 185	
GCT TAT AAG GTG GCT CCG GTC GTT TTT AGC GAG CTT GGG GCT GAT GTG	627
Ala Tyr Lys Val Ala Pro Val Val Phe Ser Glu Leu Gly Ala Asp Val	
190 195 200	

-342-

TTA GTG ATT AAT GAT GAG-CCT AAC GGG TGT AAC ATT AAT GAT CAA TGC	675
Leu Val Ile Asn Asp Glu Pro Asn Gly Cys Asn Ile Asn Asp Gln Cys	
205 210 215	
GGG GCT TTA CAC CCC AAC CAA TTA AGC CAG GAA GTG AAA AAA TAC CGC	723
Gly Ala Leu His Pro Asn Gln Leu Ser Gln Glu Val Lys Lys Tyr Arg	
220 225 230	
GCA GAT TTA GGC TTT GCT TTT GAT GGC GAT GCT GAC AGG CTA GTG GTG	771
Ala Asp Leu Gly Phe Ala Phe Asp Gly Asp Ala Asp Arg Leu Val Val	
235 240 245 250	
GTG GAT AAT TTA GGG AAT ATC GTG CAT GGG GAT AAG CTT TTA GGG GTG	819
Val Asp Asn Leu Gly Asn Ile Val His Gly Asp Lys Leu Leu Gly Val	
255 260 265	
TTA GGG GTT TAT CAA AAA TCT AAA AAC GCC CTT TCT TCT CAA GCG GTT	867
Leu Gly Val Tyr Gln Lys Ser Lys Asn Ala Leu Ser Ser Gln Ala Val	
270 275 280	
GTC GCC ACA AAC ATG AGC AAT TTA GCC CTT AAA GAA TAT TTA AAA TCC	915
Val Ala Thr Asn Met Ser Asn Leu Ala Leu Lys Glu Tyr Leu Lys Ser	
285 290 295	
CAA GAT TTG GAA TTG AAG CAT TGC GCG ATT GGG GAT AAG TTT GTG AGC	963
Gln Asp Leu Glu Leu Lys His Cys Ala Ile Gly Asp Lys Phe Val Ser	
300 305 310	
GAA TGC ATG CAA TTG AAT AAA GCC AAT TTT GGA GGC GAG CAA AGC GGG	1011
Glu Cys Met Gln Leu Asn Lys Ala Asn Phe Gly Gly Glu Gln Ser Gly	
315 320 325 330	
CAT ATC ATT TTT AGC GAT TAC GCT AAA ACA GGC GAT GGT TTG GTG TGC	1059
His Ile Ile Phe Ser Asp Tyr Ala Lys Thr Gly Asp Gly Leu Val Cys	
335 340 345	
GCT TTG CAA GTG AGC GCG TTA GTG TTA GAA AGC AAG CAG GTA AGC TCT	1107
Ala Leu Gln Val Ser Ala Leu Val Leu Glu Ser Lys Gln Val Ser Ser	
350 355 360	
GTT GCG TTA AAC CCC TTT GAA TTA TAC CCC CAA AGC CTA GTG AAT TTG	1155
Val Ala Leu Asn Pro Phe Glu Leu Tyr Pro Gln Ser Leu Val Asn Leu	
365 370 375	
AAT GTC CAA AAA AAG CCC CCT TTA GAA AGC CTG AAA GGT TAT AGC GCT	1203
Asn Val Gln Lys Lys Pro Pro Leu Glu Ser Leu Lys Gly Tyr Ser Ala	
380 385 390	
CTT TTA AAA GAA TTA GAC AAG CTA GAA ATC CGC CAT TTG ATC CGT TAT	1251
Leu Leu Lys Glu Leu Asp Lys Leu Glu Ile Arg His Leu Ile Arg Tyr	
395 400 405 410	
AGC GGC ACT GAA AAC AAA TTG CGA ATC CTT TTA GAA GCT AAA GAT GAA	1299
Ser Gly Thr Glu Asn Lys Leu Arg Ile Leu Leu Glu Ala Lys Asp Glu	
415 420 425	

-343-

AAG CTT TTA GAA TCC AAA-ATG CAA GAA TTA AAA GAG TTT TTT GAA GGG 1347
 Lys Leu Leu Glu Ser Lys Met Gln Glu Leu Lys Glu Phe Phe Glu Gly
 430 435 440

CAT TTG TGC TAAAAACCAC TAAAAAAGC CTGTTGGTTT TTATGG 1392
 His Leu Cys
 445

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 445 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

Met	Lys	Ile	Phe	Gly	Thr	Asp	Gly	Val	Arg	Gly	Lys	Ala	Gly	Val	Lys	1	5	10	15
Leu	Thr	Pro	Met	Phe	Val	Met	Arg	Leu	Gly	Ile	Ala	Ala	Gly	Leu	Tyr	20	25	30	
Phe	Lys	Lys	His	Ser	Gln	Thr	Asn	Lys	Ile	Leu	Ile	Gly	Lys	Asp	Thr	35	40	45	
Arg	Lys	Ser	Gly	Tyr	Met	Val	Glu	Asn	Ala	Leu	Val	Ser	Ala	Leu	Thr	50	55	60	
Ser	Ile	Gly	Tyr	Asn	Val	Ile	Gln	Ile	Gly	Pro	Met	Pro	Thr	Pro	Ala	65	70	75	80
Ile	Ala	Phe	Leu	Thr	Glu	Asp	Met	Arg	Cys	Asp	Ala	Gly	Ile	Met	Ile	85	90	95	
Ser	Ala	Ser	His	Asn	Pro	Phe	Glu	Asp	Asn	Gly	Ile	Lys	Phe	Phe	Asn	100	105	110	
Ser	Tyr	Gly	Tyr	Lys	Leu	Lys	Glu	Glu	Glu	Lys	Ala	Ile	Glu	Glu		115	120	125	
Ile	Phe	His	Asp	Glu	Glu	Leu	His	Ser	Ser	Tyr	Lys	Val	Gly	Glu		130	135	140	
Ser	Val	Gly	Ser	Ala	Lys	Arg	Ile	Asp	Asp	Val	Ile	Gly	Arg	Tyr	Ile	145	150	155	160
Ala	His	Leu	Lys	His	Ser	Phe	Pro	Lys	His	Leu	Asn	Leu	Gln	Ser	Leu	165	170	175	
Arg	Ile	Val	Leu	Asp	Thr	Ala	Asn	Gly	Ala	Ala	Tyr	Lys	Val	Ala	Pro	180	185	190	
Val	Val	Phe	Ser	Glu	Leu	Gly	Ala	Asp	Val	Leu	Val	Ile	Asn	Asp	Glu	195	200	205	
Pro	Asn	Gly	Cys	Asn	Ile	Asn	Asp	Gln	Cys	Gly	Ala	Leu	His	Pro	Asn	210	215	220	
Gln	Leu	Ser	Gln	Glu	Val	Lys	Lys	Tyr	Arg	Ala	Asp	Leu	Gly	Phe	Ala	225	230	235	240
Phe	Asp	Gly	Asp	Ala	Asp	Arg	Leu	Val	Val	Val	Asp	Asn	Leu	Gly	Asn	245	250	255	
Ile	Val	His	Gly	Asp	Lys	Leu	Leu	Gly	Val	Leu	Gly	Val	Tyr	Gln	Lys	260	265	270	

-344-

Ser	Lys	Asn	Ala	Leu	Ser	Ser	Gln	Ala	Val	Val	Ala	Thr	Asn	Met	Ser
		275					280					285			
Asn	Leu	Ala	Leu	Lys	Glu	Tyr	Leu	Lys	Ser	Gln	Asp	Leu	Glu	Leu	Lys
		290					295				300				
His	Cys	Ala	Ile	Gly	Asp	Lys	Phe	Val	Ser	Glu	Cys	Met	Gln	Leu	Asn
305					310					315					320
Lys	Ala	Asn	Phe	Gly	Gly	Glu	Gln	Ser	Gly	His	Ile	Ile	Phe	Ser	Asp
				325					330					335	
Tyr	Ala	Lys	Thr	Gly	Asp	Gly	Leu	Val	Ser	Cys	Ala	Leu	Gln	Val	Ser
			340					345						350	
Leu	Val	Leu	Glu	Ser	Lys	Gln	Val	Ser	Ser	Val	Ala	Leu	Asn	Pro	Phe
		355					360						365		
Glu	Leu	Tyr	Pro	Gln	Ser	Leu	Val	Asn	Leu	Asn	Val	Gln	Lys	Lys	Pro
		370				375					380				
Pro	Leu	Glu	Ser	Leu	Lys	Gly	Tyr	Ser	Ala	Leu	Leu	Lys	Glu	Leu	Asp
385					390					395					400
Lys	Leu	Glu	Ile	Arg	His	Leu	Ile	Arg	Tyr	Ser	Gly	Thr	Glu	Asn	Lys
				405					410					415	
Leu	Arg	Ile	Leu	Leu	Glu	Ala	Lys	Asp	Glu	Lys	Leu	Leu	Glu	Ser	Lys
			420					425					430		
Met	Gln	Glu	Leu	Lys	Glu	Phe	Phe	Glu	Gly	His	Leu	Cys			
		435					440					445			

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

GCCGGATCCA TGACTTATGG GSTATGGGGAA

30

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

GCCCTCGAGA CTTTATTGA TTCACCATTT CATT

34

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:

-345-

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

GCCGGATCCA TCGCTGAAGA AAATGGGGCG

30

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GCCCGGCCGC CCTAAAAACT ATAAACATAA CTC

33

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

GCCGGATCCG GTATTAGGAA GCTTATACCA TC

32

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

GCCCTCGAGA AGTTCTATTT TTAATTCCTT GAGAG

35

(2) INFORMATION FOR SEQ ID NO:177:

-346-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

GCCGGATCCT CTGATAGCCA TAAAGAAAAA AAGGAC

36

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

GCCCTCGAGA TCTTTAGAAA TCAACCCCA AAGC

34

(2) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

GCCGGATCCG ACTTAGAACA TTTTAACACG CTC

33

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

GCCCTCGAGT CATTTTAAAC GACTCAAAAC AAA

33

(2) INFORMATION FOR SEQ ID NO:181:

-347-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

GCCGGATCCG GCCAAAGCGT GCGCACTTAT TGG

33

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

GCCCTCGAGT TATTGTTCCA ACCCCCACGC ATC

33

(2) INFORMATION FOR SEQ ID NO:183:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

GCCGGATCCA AGAGCAATGC TGATGACAAA CC

32

(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

GCCCTCGAGT TATGAGTTAA AGCCCCTTGT CC

32

(2) INFORMATION FOR SEQ ID NO:185:

-348-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

GCCGGATCCG AATCAGTAAA AACAGGAAAA AC

32

(2) INFORMATION FOR SEQ ID NO:186:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

GCCCTCGAGC GGCTCTTTGG AGTTTTATTG

30

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

GCCGGATCCA TCATTCCCTC TCGCTCTATG G

31

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

GCCCTCGAGA CCTTAATGCG TTGCGTTTTC TTT

33

(2) INFORMATION FOR SEQ ID NO:189:

-349-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

GCCGAGCTCC AAGCAAAAAA ATGTCAATTA AAAGGG

36

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

GCCCTCGAGG TCTAAATTAG AATAAGTGTT GTT

33

-350-

1. An isolated polynucleotide that encodes:

(i) a polypeptide comprising an amino acid sequence that is homologous to the amino acid sequence of a *Helicobacter* polypeptide, wherein said amino acid sequence of said *Helicobacter* polypeptide is selected from the group consisting of the amino acid sequences as shown in SEQ ID NO:2 (GHPO 13), SEQ ID NO:4 (GHPO 73), SEQ ID NO:6 (GHPO 90), SEQ ID NO:8 (GHPO 107), SEQ ID NO:10 (GHPO 136), SEQ ID NO:12 (GHPO 191), SEQ ID NO:14 (GHPO 213), SEQ ID NO:16 (GHPO 240), SEQ ID NO:18 (GHPO 408), SEQ ID NO:20 (GHPO 411), SEQ ID NO:22 (GHPO 419), SEQ ID NO:24 (GHPO 431), SEQ ID NO:26 (GHPO 474), SEQ ID NO:28 (GHPO 591), SEQ ID NO:30 (GHPO 596), SEQ ID NO:32 (GHPO 699), SEQ ID NO:34 (GHPO 724), SEQ ID NO:36 (GHPO 730), SEQ ID NO:38 (GHPO 761), SEQ ID NO:40 (GHPO 804), SEQ ID NO:42 (GHPO 805), SEQ ID NO:44 (GHPO 812), SEQ ID NO:46 (GHPO 879), SEQ ID NO:48 (GHPO 888), SEQ ID NO:50 (GHPO 986), SEQ ID NO:52 (GHPO 1056), SEQ ID NO:54 (GHPO 1081), SEQ ID NO:56 (GHPO 1100), SEQ ID NO:58 (GHPO 1140), SEQ ID NO:60 (GHPO 1148), SEQ ID NO:62 (GHPO 1200), SEQ ID NO:64 (GHPO 1212), SEQ ID NO:66 (GHPO 1258), SEQ ID NO:68 (GHPO 1263), SEQ ID NO:70 (GHPO 1273), SEQ ID NO:72 (GHPO 1284), SEQ ID NO:74 (GHPO 1299), SEQ ID NO:76 (GHPO 1327), SEQ ID NO:78 (GHPO 1346), SEQ ID NO:80 (GHPO 1378), SEQ ID NO:82 (GHPO 1412), SEQ ID NO:84 (GHPO 1443), SEQ ID NO:86 (GHPO 1466), SEQ ID NO:88 (GHPO 1476), SEQ ID NO:90 (GHPO 1536), SEQ ID NO:92 (GHPO 1559), SEQ ID NO:94 (GHPO 427), SEQ ID NO:96 (GHPO 1045), SEQ ID NO:98 (GHPO 1262), SEQ ID NO:100 (GHPO 1688), SEQ ID NO:102 (GHPO 1538), SEQ ID NO:104 (GHPO 346), SEQ ID NO:106 (GHPO 1012), SEQ ID NO:108 (GHPO 470), SEQ ID NO:110 (GHPO 1398), SEQ ID NO:112 (GHPO 1550), SEQ ID NO:114 (GHPO 276), SEQ ID NO:116 (GHPO 1501), SEQ ID NO:118 (GHPO 706), SEQ ID NO:120 (GHPO 1001), SEQ ID NO:122 (GHPO 732), SEQ ID NO:124 (GHPO 329), SEQ ID NO:126 (GHPO 574),

-351-

SEQ ID NO:128 (GHPO 1190), SEQ ID NO:130 (GHPO 1374), SEQ ID NO:132 (GHPO 1620), SEQ ID NO:134 (GHPO 956), SEQ ID NO:136 (HPO 98), SEQ ID NO:138 (GHPO 689), SEQ ID NO:140 (GHPO 208), SEQ ID NO:142 (GHPO 296), SEQ ID NO:144 (GHPO 726), SEQ ID NO:146 (GHPO 1026), SEQ ID NO:148 (GHPO 1301), SEQ ID NO:150 (GHPO 1536), SEQ ID NO:152 (GHPO 166), SEQ ID NO:154 (GHPO 253), SEQ ID NO:156 (GHPO 297), SEQ ID NO:158 (GHPO 615), SEQ ID NO:160 (GHPO 1278), SEQ ID NO:162 (GHPO 1282), SEQ ID NO:164 (GHPO 1420), SEQ ID NO:166 (GHPO 1484), SEQ ID NO:168 (GHPO 1719), and SEQ ID NO:170 (GHPO 1252); or

(ii) a derivative of said polypeptide encoded by said polynucleotide.

2. The isolated polynucleotide of claim 1, which encodes a mature form of said polypeptide.

3. The isolated polynucleotide of claim 1 or 2, wherein the polynucleotide is a DNA molecule.

4. A compound, in a substantially purified form, that is the mature form or a derivative of a polypeptide comprising an amino acid sequence that is homologous to a *Helicobacter* amino acid sequence that is selected from the group consisting of the amino acid sequences as shown in SEQ ID NO:2 (GHPO 13), SEQ ID NO:4 (GHPO 73), SEQ ID NO:6 (GHPO 90), SEQ ID NO:8 (GHPO 107), SEQ ID NO:10 (GHPO 136), SEQ ID NO:12 (GHPO 191), SEQ ID NO:14 (GHPO 213), SEQ ID NO:16 (GHPO 240), SEQ ID NO:18 (GHPO 408), SEQ ID NO:20 (GHPO 411), SEQ ID NO:22 (GHPO 419), SEQ ID NO:24 (GHPO 431), SEQ ID NO:26 (GHPO 474), SEQ ID NO:28 (GHPO 591), SEQ ID NO:30 (GHPO 596), SEQ ID NO:32 (GHPO 699), SEQ ID NO:34 (GHPO 724), SEQ ID NO:36 (GHPO 730), SEQ ID NO:38 (GHPO 761), SEQ ID NO:40 (GHPO 804), SEQ ID NO:42 (GHPO 805), SEQ ID

-352-

NO:44 (GHPO 812), SEQ ID NO:46 (GHPO 879), SEQ ID NO:48 (GHPO 888),
SEQ ID NO:50 (GHPO 986), SEQ ID NO:52 (GHPO 1056), SEQ ID NO:54 (GHPO
1081), SEQ ID NO:56 (GHPO 1100), SEQ ID NO:58 (GHPO 1140), SEQ ID NO:60
(GHPO 1148), SEQ ID NO:62 (GHPO 1200), SEQ ID NO:64 (GHPO 1212), SEQ
5 ID NO:66 (GHPO 1258), SEQ ID NO:68 (GHPO 1263), SEQ ID NO:70 (GHPO
1273), SEQ ID NO:72 (GHPO 1284), SEQ ID NO:74 (GHPO 1299), SEQ ID NO:76
(GHPO 1327), SEQ ID NO:78 (GHPO 1346), SEQ ID NO:80 (GHPO 1378), SEQ
ID NO:82 (GHPO 1412), SEQ ID NO:84 (GHPO 1443), SEQ ID NO:86 (GHPO
1466), SEQ ID NO:88 (GHPO 1476), SEQ ID NO:90 (GHPO 1536), SEQ ID NO:92
10 (GHPO 1559), SEQ ID NO:94 (GHPO 427), SEQ ID NO:96 (GHPO 1045), SEQ ID
NO:98 (GHPO 1262), SEQ ID NO:100 (GHPO 1688), SEQ ID NO:102 (GHPO
1538), SEQ ID NO:104 (GHPO 346), SEQ ID NO:106 (GHPO 1012), SEQ ID
NO:108 (GHPO 470), SEQ ID NO:110 (GHPO 1398), SEQ ID NO:112 (GHPO
1550), SEQ ID NO:114 (GHPO 276), SEQ ID NO:116 (GHPO 1501), SEQ ID
15 NO:118 (GHPO 706), SEQ ID NO:120 (GHPO 1001), SEQ ID NO:122 (GHPO
732), SEQ ID NO:124 (GHPO 329), SEQ ID NO:126 (GHPO 574), SEQ ID NO:128
(GHPO 1190), SEQ ID NO:130 (GHPO 1374), SEQ ID NO:132 (GHPO 1620), SEQ
ID NO:134 (GHPO 956), SEQ ID NO:136 (HPO 98), SEQ ID NO:138 (GHPO 689),
SEQ ID NO:140 (GHPO 208), SEQ ID NO:142 (GHPO 296), SEQ ID NO:144
20 (GHPO 726), SEQ ID NO:146 (GHPO 1026), SEQ ID NO:148 (GHPO 1301), SEQ
ID NO:150 (GHPO 1536), SEQ ID NO:152 (GHPO 166), SEQ ID NO:154 (GHPO
253), SEQ ID NO:156 (GHPO 297), SEQ ID NO:158 (GHPO 615), SEQ ID NO:160
(GHPO 1278), SEQ ID NO:162 (GHPO 1282), SEQ ID NO:164 (GHPO 1420), SEQ
ID NO:166 (GHPO 1484), SEQ ID NO:168 (GHPO 1719), and SEQ ID NO:170
25 (GHPO 1252); or

(ii) a derivative of said polypeptide.

-353-

5. A method of preventing or treating *Helicobacter* infection in a mammal, said method comprising administering to said mammal a prophylactically or therapeutically effective amount of a compound of claim 4.

6. The method of claim 5, further comprising administering to said mammal an antibiotic, an antisecretory agent, a bismuth salt, or a combination thereof.

7. The method of claim 6, wherein said antibiotic is selected from the group consisting of amoxicillin, clarithromycin, tetracycline, metronidazole, and erythromycin, and said bismuth salt is selected from the group consisting of bismuth subcitrate and bismuth subsalicylate.

8. The method of claim 6, wherein said antisecretory agent is a proton pump inhibitor, an H₂-receptor antagonist, or a prostaglandin analog.

9. The method of claim 8, wherein said proton pump inhibitor is selected from the group consisting of omeprazole, lansoprazole, and pantoprazole; said H₂-receptor antagonist is selected from the group consisting of ranitidine, cimetidine, famotidine, nizatidine, and roxatidine; and said prostaglandin analog is selected from the group consisting of misoprostil and enprostil.

10. The method of claim 5, further comprising administering to said mammal a prophylactically or therapeutically effective amount of a second *Helicobacter* polypeptide or a derivative thereof.

11. The method of claim 10, wherein the second *Helicobacter* polypeptide is a *Helicobacter* urease, a subunit, or a derivative thereof.

-354-

12. A composition comprising a compound of claim 4, together with a physiologically acceptable diluent or carrier.

13. The composition of claim 12, further comprising an adjuvant.

14. The composition of claim 12, further comprising a second *Helicobacter* polypeptide or a derivative thereof.

15. The composition of claim 14, wherein said second *Helicobacter* polypeptide is a *Helicobacter* urease, a subunit, or a derivative thereof.

16. A method of preventing or treating *Helicobacter* infection in a mammal, said method comprising administering to said mammal a prophylactically or therapeutically effective amount of a polynucleotide of claim 1.

17. A composition comprising a viral vector, in the genome of which is inserted a DNA molecule of claim 1, said DNA molecule being placed under conditions for expression in a mammalian cell and said viral vector being admixed with a physiologically acceptable diluent or carrier.

18. A composition that comprises a bacterial vector comprising a DNA molecule of claim 1, said DNA molecule being placed under conditions for expression and said bacterial vector being admixed with a physiologically acceptable diluent or carrier.

19. The composition of claim 18, wherein said vector is selected from the group consisting of *Shigella*, *Salmonella*, *Vibrio cholerae*, *Lactobacillus*, Bacille bilié de Calmette-Guérin, and *Streptococcus*.

-355-

20. A composition comprising a polynucleotide of claim 1, together with a physiologically acceptable diluent or carrier.

21. The composition of claim 20, wherein said polynucleotide is a DNA molecule that is inserted in a plasmid that is unable to replicate and to substantially integrate in a mammalian genome and is placed under conditions for expression in a mammalian cell.

22. An expression cassette comprising a DNA molecule of claim 1, said DNA molecule being placed under conditions for expression in a procaryotic or eucaryotic cell.

23. A process for producing a compound of claim 4, which comprises culturing a procaryotic or eucaryotic cell transformed or transfected with an expression cassette of claim 22, and recovering said compound from the cell culture.

24. A method of preventing or treating *Helicobacter* infection in a mammal, said method comprising administering to said mammal a prophylactically or therapeutically effective amount of an antibody that binds to the compound of claim 4.

1 / 1

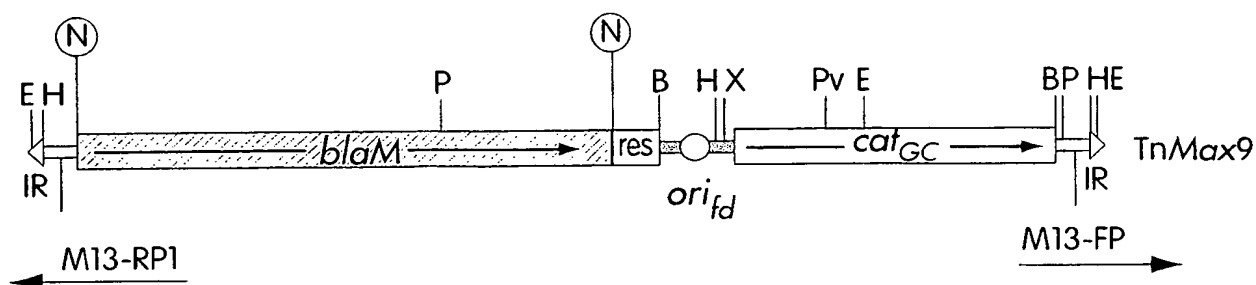


Fig.1A

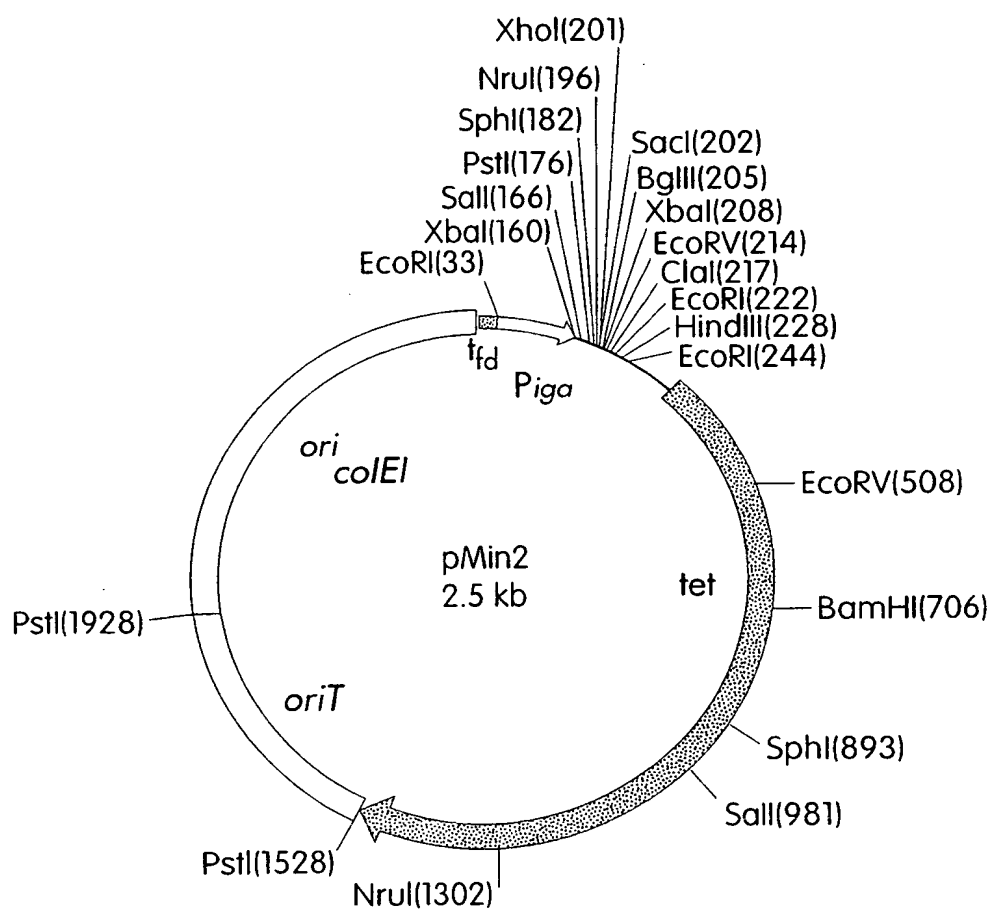


Fig.1B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/21353

A. CLASSIFICATION F SUBJECT MATTER

IPC(6) : CO7H 21/02; CO7K 1/00; A61K 39/02; AO1N 37/18
US CL : 536/23.1; 530/350; 424/190.1; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 530/350; 424/190.1; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Geneseq, Pir, Swiss Prot, Business & Industry, Health News Daily, Biosis, Medline, Int. Pharm Abst, Derwent, Scisearch, Embase, Pascal, Gen Sci Abst, Diss Abst

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 91/09049 A1 (RESEARCH EXPLOITATION LIMITED) 27 June 1991, see entire document	1-4 and 12-14
X	WO 96/33732 A1 (ORAVAX, INC) 31 October 1996, see entire document	1-16 and 24
X,P	WO 96/40893 A1 (ASTRA AKTIEBOLAG) 19 December 1996, see abstract and claims.	1-16 and 24
X,P	WO 97/37044 A1 (ASTRA AKTIEBOLAG) 09 October 1997, see abstract and claims.	1-16 and 24
Y	Monath, T. P. et al., Development of Recombinant Helicobacter pylori Urease as an Oral Vaccine: Current Status. GUT, 1995. 37 (SUPP 1)A52.	10, 11, 15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

B earlier document published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

A

document member of the same patent family

Date of the actual completion of the international search

29 MARCH 1998

Date of mailing of the international search report

20 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile N (703) 305-3230

Authorized officer

SUSAN A. LOBNO

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/21353

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 97/19098 A1 (ASTRA AKTIEBOLAG) 29 May 1997, see entire document.	1-16 and 24
Y	CLAYTON, C.L. et al, Sensitive Detection of Helicobacter pylori by Using Polymerase Chain Reaction. J Clin Microbiol, January 1992, Vol 30. No. 1. pages 192-200, see entire document.	1-4 and 12-14
Y,P	TOMB, JEAN-F. et al. The Complete Genome Sequence of the Gastric Pathogen Helicobacter pylori. Nature. 07 August 1997, Vol 388. No. 6642. Pages 539-547, see entire document.	1-4 and 12-14
Y, P	MYERS, G. et al. Oral Immunization with Recombinant Urease Confers Long-Lasting Immunity. GUT, 1996. 39(Suppl 2), A94, see abstract.	10-11 and 15
Y	THIJS, J. C. et al, Efficacy and Side Effects of a Triple Drug Regimen for the Eradication of Helicobacter pylori. Scand J Gastroenterol. 1993. Vol 28. Pages 934-938.	6-9 and 12-14
Y	GRAHAM, DAVID Y et al. Effect of Triple Therapy (Antibiotics plus Bismuth) on Duodenal Ulcer Healing. Annals of Internal Medicine. 15 August 1991. Vol 115. No. 4. Pages 266-269, see entire document.	6-9
Y	NAKAO, M et al. Antibacterial Properties of Lansoprazole Alone and in Combination with Antimicrobial Agents against Helicobacter pylori. Eur J Clin Microbiol Infect Dis, 1995. Vol 14. No. 5, pages 391-399, see entire document.	6-9
Y	BORODY, T J et al. Omeprazole Enhances Efficacy of Triple Therapy in Eradicating Helicobacter pylori. GUT. 1995. Vol 37. Pages 477-481, see entire document.	6-9

Form PCT/ISA/210 (continuation of second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/21353

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-16 and 24

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/21353

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

1. This International Search Authority has found 22 inventions claimed in the International Application covered by the claims indicated below:

Group I, claim(s) 1-16 and 24, drawn to isolated polynucleotide that encodes a polypeptide comprising an amino acid sequence; a composition comprising a compound that is the mature form of said polypeptide; and a method of preventing or treating *Helicobacter* infection in a mammal.

Group II, claim(s) 17-21, drawn to a composition comprising a viral or bacterial vector.

Group III, claim(s) 22-23, drawn to an expression cassette and a process for producing a compound that is the mature form of said polypeptide which comprises culturing a procaryotic or eucaryotic cell transformed or transfected with said expression cassette.

The claims are deemed to correspond to the species listed above in the following manner:

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Group I, claims 1 and 4 contain 85 structurally different sequence species. The first 10 sequence species will be included with the search of Group I. Each subsequent species election will consist of 4 sequences (possible 19 additional species).

and it considers that the International Application does not comply with the requirements of unity of invention (Rules 13.1, 13.2 and 13.3) for the reasons indicated below:

The inventions listed as Groups I, II, and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the products and method of treating of Group I contain numerous different species. The products of Groups II and III do not require a single specie of Group I and the product of Group I can be used in other products or methods than those of Groups II or III, such as using the polypeptide as an immunogen for producing antibodies or in a diagnostic immunoassay.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: each of the claimed sequences are structurally different.

If no additional fees are paid, Group I, SEQ ID NOs 2, 4, 6, 8, 10, 12, 14, 16, 18 and 20 will be searched.

